

Qy	2211	ggagagggaaaaacagaagctccaagcctctgaaatgatttgtagcaagcgccgtgaa	2270
Db	2487	GGAGAGGAGGAATACAGAGCTACAGCCCTCTGAAACGATTTGCTGACAAACGCCCTCAA	2546
Qy	2271	gctcgattatgaagaattactccctgctctaaagaagtaactacagctgagggaagaat	2330
Db	2547	GCTTGACTATGAAGAANTCAGTCGCTGCTTTAAAGAAGCTACTACAGTGTGGGAAAAGAT	2606
Qy	2331	gcttagcactccaggagatcaaaaattaaagtttgacatggaaaaaattgcactcgctgt	2390
Db	2607	GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAGTGCACTCAGCTGT	2666
Qy	2391	tgggcaaggtgtgccacgtcatccacggaggtagaatctcgaaattcttagctgagcaatt	2450
Db	2667	TGGCAAGGTGTGCCACGTCAACCGAGGTGAGATCTGGAATTTCTAGCTGAGCAGTT	2726
Qy	2451	ccacctaaacaccagttccccagcaaacacagaccaagatgcccatacaaaagaact	2510
Db	2727	CCACCTTAACACCCATTCTCTAGTAACAGCAGGCCAAGAGCGTGCCCTACAAAGAGCT	2786
Qy	2511	cttaagcagctgactcccacgacgacatgcgattcttattgaccttgggcgaacctctcc	2570
Db	2787	CCTGAAGAAGCTGACCTCGCACGACGACGCCATTCTCATCGACCTCGGGCGAACCTTTC	2846
Qy	2571	tacacaccatactctctgcccagcttggagcaggacagctatcgottttacaacatttt	2630
Db	2847	AACACATCCATACTTCTCTGCCAGCTTGGAGCAGGTGCTGCTACTTTACAACATCT	2906
Qy	2631	gaaggcctactcactctagacacagaagtagggatatatgcacaaagttctcagctttagc	2690
Db	2907	GAAGGCTTACTCGCTTCTGGACGAGAGGTGGATCTGCAAGGTCTCAGCTTTGTGGC	2966
Qy	2691	aggcatttgccttctcatatagtgaggaaagagcgctttaaaatgctcaagttctcgt	2750
Db	2967	AGCATTTGCTCTTCACATGAGTGAGGAAGAGCGCTTCAAGATGCTCAAGTTCCTGAT	3026
Qy	2751	gtttgacatggggctcggaacagta tcggccagacatgattatttcagatccagat	2810
Db	3027	GTTTGATAGTGGGCTCGGAAACAGTATCGGCCAGACATGATTATTTTGAGATCCAGAT	3086
Qy	2811	gtaccagctctggaggtgcttcattgattaccacagaagacctacaaatcacctggagga	2870
Db	3087	GTACCAGCTGTACAGGCTCCTCTCAGCATTTACCACGGAGACCTCTACAAACACCTGGGAAGA	3146
Qy	2871	gcacgagatcggccccagcctcagctgcgccccctggctcctcaccatgtttgcctcaca	2930
Db	3147	GCAGGAGACTGGCCCCCTACGTAGCGGGCTCCCTGGTTTCTCACCGTGTTCGCTCACA	3206
Qy	2931	gttcocgctgggattcgttagccagagcttcttgatatgattttcttcagggaacagaggt	2990
Db	3207	GTTTCCCACTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTTCTCTCAGGGATCAGAGGT	3266
Qy	2991	cataittaaaggctttaagtctgttgggaagccaataagcccttgattctcgaagcatga	3050
Db	3267	CATATTAAAGTAGCTTTAAGTCTTTTGGGAGGCCATAAGCCCTTGATTCTACAGCATGA	3326
Qy	3051	aaacctagaaaccatagttgactttataaaagcacgctacccaaaccttggtcttggtaca	3110
Db	3327	GAACCTGGAAACCATCGTGGACTTCATTAAGNACACACTCTCCCAACCTGGGCCCTGGTGCA	3386
Qy	3111	gatggaaaagaccatcaatcaggtattgaaatggacatcgctaaacagttacaagctta	3170
Db	3387	GATGGAGAAGACCATCAGTCAGGTGTTTTGAGATGGACATCGCCAAAGCAGCTCCAGGCCTA	3446
Qy	3171	tgaagttgagtaccagctcctcaagaagaacttatcgattcctctcctctcagtagcaaa	3230
Db	3447	TGAGGTCGAGTACCAGCTCGTCCAGGAGAGCTTATTAGTCTCTCGGCTCTCAGTGCACAA	3506
Qy	3231	ccaaagaatggataaattagcagaaaccaaacagcagcttacgcgaacagaaaccttgacct	3290
Db	3507	CCAAAGAATGGAGAAATTTGGAGAAANACCAACACAGACGCTTGGCGAAACAGAACCTTGACCT	3566
Qy	3291	ccttgacagttgcaggtggcaaatggtaggatccaaagccttgaggccaccattgagaa	3350

```

Db    3567 CCTGGAGCAGTTCCAGGTGCGCAAAATCTAGSATTCCAAGACCTTGAAGCCACGGTAGAGAA   3626
Qy    3351 gctccgtgacagtgagcagaactgaagcaggcccatgcttacccttagaaactggagcggtc   3410
Db    3627 ACTTCATTACCAGGCAGAGTAAGCTGAAGCAGCGTGCGTGTACCTCTGGAGGTGGAGCG- -T   3684
Qy    3411 ggccctgtcacagcgttgagagagtgcggcgggagcgcagagagccccagcagccgccggga   3470
Db    3685 CGCCTGTCTCATGTGTGGAGGAGCTGCGGAGGCCAAGCGCCGCCGCCACTCCAGA   3744
Qy    3471 gcttgagtgcacgcagcccagccccagcggcgacctgacagct   3512
Db    3745 GCCAGACTGCACCACGCTGAGGCCACAGCGCATGTGACCGCT   3786

RESULT      2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEX: (703)683-4109
; FAX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; CLONE: ptzqpt-fls
; IMMEDIATE SOURCE:
; ORGANISM: Gallus gallus
; PROJECT:
; CONTRACT grant sponsor:
; CONTRACT number:
; INVESTIGATOR:
; INSTITUTION:
; AVAILABILITY STATEMENT:
; ORDERING INFORMATION:
; STOCK NO.:
; DISTRIBUTION STATEMENT:
; FIELD OF SEARCH:
; ABSTRACT:

```

Query Match	1.6%;	Score 57.8;	DB 1;	Length 7218;
Best Local Similarity	2.9%;	Pred. No. 9.7e-06;		


```

; LENGTH: 9515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas aeruginosa
; STRAIN: Clinical Isolate P3-2
; US-08-920-827-13

```

Query Match	1.3%	Score 45.4;	DB 1;	Length 9515;
Best Local Similarity	46.2%;	Pred. No. 0.03%;		
Matches 151; Conservative	0;	Mismatches 176;	Indels	Gaps
Qy 438	catccgtcaggcggggaagatcgcccggaagagagctgcactgcacgtcccggtccgagttcga	497		
Db				
Qy 802	CATGGCAAGCCGGTGATGGACGGCTTGGACATCGATGTACCGCGCGCCACAGTCTT	861		
Db				
Qy 498	cgacacgttttccaaagattcgagggtctcttctggccgcgctgacggtggcgcacaa	557		
Db				
Qy 862	CGCCTGGTATCGGAAAGCCTCGACAAGCTTACGACCAGGTGCGCGCGCGGCCACGA	921		
Db				
Qy 558	gaaggctccgcgggcctgatcgacgagtgcatcgagaagttcaatcacgtcagcgcag	617		
Db				
Qy 922	GACCTGGCCACCATTTACCCGCGTCCGCTGGGGGTGATCGCGCGGTGTCGGTGAA	981		
Db				
Qy 618	ccgggggtccgagagcccccgccccaccgccccatgctcgccgccacagaggagccaga	677		
Db				
Qy 982	CTTCCCGCTGCACATGCGCGCTGGAAAGCTCGCCCCCGCCCTGGCGCGCGCAACTCGGT	1041		
Db				
Qy 678	gcttgtgcgaagccccatggcgaagtctcttccacagcccgctgctgcctggtgacct	737		
Db				
Qy 1042	GGTGCCTAAGCCGCGCAGAGAGTGCCTGTTCTCCGCCCTGGCCCGGAGCTGGGCCCT	1101		
Db				
Qy 738	taggaagagctgcagatggggacct	764		
Db				
Qy 1102	GGAGCGGGGTGCGGAAAGCGTCT	1128		
Db				

```

RESULT 5
US-08-921-177-13
; Sequence 13, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547

```

```

, REFERENCE/DOCKET NUMBER: 19036/32420
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 312/474-6300
,
, TELEFAX: 312/474-0448
,
, TELEX: 25-3856
,
, INFORMATION FOR SEQ ID NO: 13:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 9515 base pairs
,
, TYPE: nucleic acid
,
, STRANDEDNESS: double
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: Genomic DNA
,
, ORIGINAL SOURCE:
,
, ORGANISM: Pseudomonas aeruginosa
,
, STRAIN: Clinical Isolate P2-2
,
, US-08-921-177-13

```

[illegible]

```

RESULT      6
US-08-362-577C-13
; Sequence 13, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Ege, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/362,577C
; FILING DATE: 27-MAR-1995

```


ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952-US.P2
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-997-467-2

Query Match 1.1%; Score 40.8; DB 4; Length 1030;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 102; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
Qy 440 tccgtcagcggggaagatccgcgcagaggaagtcgactgcccgcgtccgagttcgacg 499
Db 203 TCACACAGCGCGCTCTGTCGCCGCGGAGACCTCCCTGTTCCGGCTGTCGAGGCCACG 262
Qy 500 acacgtttccaagaagttcgaggtgctctctcgcgcgcgcgtgacggtggcgcacaaga 559
Db 263 GCCTCGTCCCCGACTACCTCATCGGCCACTCCATCGCGGAAGTGACCGCGGCCACCTGG 322
Qy 560 aggtctccgcgcgcgcgtgacgagtgcatcgagaagtccaatcacgtcagcgagcacc 619
Db 323 CCGGGGTCTCGATCTGCGGACGGGTGCTGCTGTCGCCACCGCGCGCGCTGATGC 382
Qy 620 gggggtcgcagagccccgcgcaca 643
Db 383 AGTCGGCCCCGGCGCGCGCGCA 406

RESULT 13
US-08-814-052-19
; Sequence 19, Application US/08814052
; Patent No. 6015783
; GENERAL INFORMATION:
; APPLICANT: von der Osten, Claus
; APPLICANT: Cherry, Joel R.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Wind, Jesper
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
; TITLE OF INVENTION: OR STAINS FROM CELLULOSIC FABRIC
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60157830 No. 6015783disk of No. 6015783th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,052
; FILING DATE: 06-MAR-1997
; CLASSIFICATION: 510
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4684.204-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-867-0123

TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2249 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-814-052-19

Query Match 1.1%; Score 40.6; DB 3; Length 2249;
Best Local Similarity 48.9%; Pred. No. 0.31;
Matches 109; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
Qy 474 gctgcactgcccgcgtccgagttcgacgacagcttttccaagaagttcgaggtctctctg 533
Db 1234 GCTGTGCGCGGCTCCAGGAATCCCTACCCGCGCCCATCTTCCACTACGCGCGGCCCC 1293
Qy 534 cggccgcgtgacggtggtcgcaagaagctccgcgcgcgcctgacgcagagtgcatcga 593
Db 1294 CGGCGCGCGCCCGGACGAGGCGCAAGCCCGGTCCGACCACTGCTGGACCTCCC 1353
Qy 594 gaagtccaatcacgtcagc 653
Db 1354 CAACCTCAAGCCGCTGTCGCGCGCGACGTGCCCTGAGCGCTTCGCAAGCGCGCGA 1413
Qy 654 tgc 696
Db 1414 CAACACGCTCGACGTCACCTCGACACACGCGGCGACGCCCTG 1456

RESULT 14
US-08-814-052-17
; Sequence 17, Application US/08814052
; Patent No. 6015783
; GENERAL INFORMATION:
; APPLICANT: von der Osten, Claus
; APPLICANT: Cherry, Joel R.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Wind, Jesper
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
; TITLE OF INVENTION: OR STAINS FROM CELLULOSIC FABRIC
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60157830 No. 6015783disk of No. 6015783th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/814,052
; FILING DATE: 06-MAR-1997
; CLASSIFICATION: 510
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4684.204-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2279 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
US-08-814-052-17

Query Match

Best Local Similarity 1.1%; Score 40.6; DB 3; Length 2279;
Matches 109; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
QY 474 gctgcaactgcccgtccgagttcgagacacacgctttttccaaagaagttcgaggtgctttctg 533
DB 1340 GCTCTCGCGCGGTCCAGGAATCCCTACCGCGCGCCATCTTCCACTACGCGCGCCCC 1399
QY 534 cgcccgctgacggtgctgcacaaagagctccgcgcgcgcctgatcgagagtgcatcg 593
DB 1400 CGCGGCGCCGCCACGACGAGGCGCAAGGCCCGGTGACACCAACTGCTGGACCTCCC 1459
QY 594 gaagttcaatcacgtcagcgccgagcgccggttcgagagcccccccaaccgccccca 653
DB 1460 CAACCTCAAGCCGCTCGTGGCGCGACGTGCCCTGAGCGGCTTCCGCAAGCGGCCGA 1519
QY 654 tgccggccccacagggagcagagcctgtgagcagggcccatg 696
DB 1520 CAACACGCTCGACGTACCTCGACACACGCGGCGCGCCCTG 1562

RESULT 15

US-08-814-052-18
; Sequence 16, Application US/08814052
; Patent No. 6015783
; GENERAL INFORMATION:
; APPLICANT: von der Osten, Claus
; APPLICANT: Cherry, Joel R.
; APPLICANT: Bjornvad, Mads E.
; APPLICANT: Vind, Jesper
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: PROCESS FOR REMOVAL OR BLEACHING OF SOILING
; TITLE OF INVENTION: OR STAINS FROM CELLULOSE FABRIC
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 60157830 No. 6015783disk of No. 6015783th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 06-MAR-1997
; CLASSIFICATION: 510
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4684.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2300 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-814-052-18

Query Match

1.1%; Score 40.6; DB 3; Length 2300;

Best Local Similarity 48.9%; Pred. No. 0.32;
Matches 109; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
QY 474 gctgcaactgcccgtccgagttcgagacacacgctttttccaaagaagttcgaggtgctttctg 533
DB 1234 GCTCTCGCGCGGTCCAGGAATCCCTACCGCGCGCCATCTTCCACTACGCGCGCCCC 1293
QY 534 cgcccgctgacggtgctgcacaaagagctccgcgcgcgcctgatcgagagtgcatcg 593
DB 1294 CGCGGCGCCGCCACGACGAGGCGCAAGGCCCGGTGACACCAACTGCTGGACCTCCC 1353
QY 594 gaagttcaatcacgtcagcgccgagcgccggttcgagagcccccccaaccgccccca 653
DB 1354 CAACCTCAAGCCGCTCGTGGCGCGACGTGCCCTGAGCGGCTTCCGCAAGCGGCCGA 1413
QY 654 tgccggccccacagggagcagagcctgtgagcagggcccatg 696
DB 1414 CAACACGCTCGACGTACCTCGACACACGCGGCGCGCCCTG 1456
Search completed: September 9, 2002, 19:07:56
Job time: 16984 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September .9, 2002, 16:41:06 : Search time 5191.84 Seconds
(without alignments)
9241.748 Million cell updates/sec

Title: US-09-762-311-4_COPY_176_3730

Perfect score: 3555

Sequence: 1 atggaacaaataacattcac.....tccacactgtccagccctt 3555

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_gss.*
- 13: em_gss_hum.*
- 14: em_gss_inv.*
- 15: em_gss_pln.*
- 16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
c					
1	833.6	23.4	1001	10	278359 HS278359 Hu
2	826.6	23.3	1051	10	BM479578 AGENCOURT
3	819.2	23.0	1007	10	BM016445 603641252
4	782.4	22.0	784	10	BG28092 602753463
5	741.8	20.9	835	10	BG759286 602710975
6	738.8	20.8	800	10	BG746377 602703644
7	737.4	20.7	1121	10	BM476629 AGENCOURT
8	729.8	20.5	1109	10	BM460573 AGENCOURT
9	722.2	20.3	952	10	BE299948 600944494
10	713.8	20.1	838	10	BI907513 603065640
11	705	19.8	807	10	BG120498 602346815
12	701	19.7	832	10	BM014940 603640877
13	689.8	19.4	851	9	AL570425 AL570425
14	683.4	19.2	754	10	BE300005 600944594
15	682.6	19.2	892	10	BE561889 601346094
16	678.6	19.1	686	10	BG769758 602744658
17	675	19.0	1108	10	BG326728 602425604

18	671.8	18.9	978	10	BG029927
19	669.6	18.8	670	9	AL557623
c	669.4	18.8	857	9	AL544177
20	668.4	18.8	750	10	BI084159
21	668.4	18.7	795	10	BE560184
22	664.6	18.6	661	9	AL046787
23	661	18.6	661	9	AL046787
24	661	18.6	948	10	BG258116
25	659.2	18.5	672	10	BE675153
26	652.8	18.4	806	9	AW043925
27	648.8	18.3	665	9	AW173375
28	647.8	18.2	812	10	BI758489
29	641	18.0	1161	10	BM466841
c	637.8	17.9	773	10	BE561312
30	636.4	17.9	862	10	BI910369
31	635.4	17.9	660	10	BE276812
32	632.4	17.8	761	10	BF797592
33	632	17.8	753	10	BF305442
34	632	17.7	722	10	BE560392
35	628	17.7	748	10	BG424359
36	628	17.7	748	10	BG424359
37	624.4	17.6	626	9	AW772402
38	621.2	17.5	626	9	AW239183
39	612	17.2	623	10	BE884440
40	604.4	17.0	649	10	BE513375
41	597	16.8	605	9	AW732976
42	596.2	16.8	622	10	BE514108
43	592.4	16.7	774	10	BI763670
44	590.2	16.6	952	10	BG424757
45	589.4	16.6	758	10	BG166636

ALIGNMENTS

RESULT 1

278359/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.

Query Match	23.43;	Score 833.6;	DB 10;	Length 1001;
Best Local Similarity	93.59;	Pred. No. 2.2e-197;		
Matches 937; Conservative	0;	Mismatches 56;	Indels	Gaps 8;
Qy 2146	acatctgtgagctccagagcgtg	gcacaaagcctattcttcaacagatactgctgtt	2205	
Db 1001	ACATCTCGTGAGCTCCGAGAGCT	GTGCCTAAGGCTATTCTTCAACAGATNCTGCTGCTN	942	
Qy 2206	agaatggagaaggaaaaatcaga	agctcctcaagctctgaaaaatgatttctgtaacaaagcgc	2265	
Db 941	AGAAATGGAGAGGAAAAATCA	GAAGCTCCAAAGCCTCTGAAAAATGATTTTCTCTGAACAAGCGC	882	
Qy 2266	ctgaagctcgattatgaagaat	tactccctgctcttaagaagtaactacagttgtggaaa	2325	
Db 881	CTGAAGCTCGAATTTATGA	AGAAATNACTTCCCTGTCTTAAGAAGTAACTTACAGTGTGGGAA	822	
Qy 2326	aagatgcttagcactccaggaag	at--caaaaaataagcttgacatgaaaaaatgcact	2383	
Db 821	AAGATGCTTAGCACTCCAG	GAAGATCCCANAAATTAAGTTTGACATGCAAAAAATGCAC	762	
Qy 2384	cggctgttgggc-aaggtgtc	acagctcatcaccgaggtgaaatctggaaa-tttctagc	2441	
Db 761	CGACTGTGTGGGCTAAGGT	GTGCCAGCTCATCACCAGAGTGAAATCTGGAATTTTCTAGC	702	
Qy 2442	tgagcaa-ttcacaccttaac	acacagctttcccagcaaacacagcgcgaagatgtgccat	2500	
Db 701	TGAGCAATTTCCACCGT	AAACACCAAGTTTCCAGCANACAGCAGCCANAGGATGTGCCAT	642	
Qy 2501	acaaagaactcttaagcagct	gactctcccagcagcatgcgattcttattgaccttgggc	2560	
Db 641	ACAAAGAACTCTTNAAG	CAGCTGACTTCCAGCAGCATGCGATTCTTATTGACCTTGGNC	582	
Qy 2561	gaaccttctcacacaccata	ctctctgccagcttggagcagacgctatcgcttt	2620	
Db 581	GAACCTTTCCACACACCC	ATACTTCTCTGCCAGCTTGGAGCAGGACGATATGCGCTT	522	
Qy 2621	acaacattttgaagcctact	cactctcttagaccaggaagtgggatttgcgaagctctca	2680	
Db 521	ACAACATTTTGAAGGCC	TACTCACTTCTAGACCAGGAAGTGGATATGGCCAAAGCTCTCA	462	
Qy 2681	gctttgtagaggcatattgt	ctctctctatgtagtgaagaagcgtttaaaaatgctca	2740	
Db 461	GCTTTGTACCAAGGCA	TTTTGGCTCTTCATATGATGTAGGAAGAGCGGTGAAGATGCTCA	402	
Qy 2741	agttctgattgtgacatggg	ctgcggaaacagtatcggccagacatgattatttac	2800	
Db 401	AGTTCTTGATG-TTGAC	ATGGGCTTNNGGNACAGNATCGGCCACACATGATTTATTTAC	343	
Qy 2801	agatccagatgataccagct	ctctcgaggttgccttcattgataccaaagacacctcaaatc	2860	
Db 342	AGANCCAGATGATACC	ANCTCTCGGGGTTTNCNTNATGATTACCACAGAGACCTCAACAATC	283	
Qy 2861	acctggagagcagagatcg	ggcccccagcctctacgctg-ccccctgggtctctcaccatg	2919	
Db 282	ACCTGGAGGAGNCAGAT	CGGCCCCAGCCCTACGCTGCCCCCTGGTTCCTCACCATG	223	
Qy 2920	tttgctcacagttccccgc	tgggattcgttagccagagcttttgatatattttctcaag	2979	
Db 222	TTTGGCTCACAGTTCC	CGCCGTGGGATTCGTATNCAAGAGNCTTTGANATGATTTTCTTCAG	163	
Qy 2980	ggaacagaggtcatatt	aaagtggcttgaagtcgttgggaagcctaagcccttgatt	3039	
Db 162	GGAACAGAGGTCA	TATTTTAAAGTGCGGNAAGNCTGTNGGGAANCCATTAAGCCNGTGATT	103	
Qy 3040	ctgcagcatgaaaacctaga	aaa-ccatagttgactttataaaaaagcagctaccacaact	3098	
Db 102	CTGCANCATGAAACCT	AGAAAGCCAGAGTNGACTTTATGTAAAGCAGCCTACCCAACCT	43	
Qy 3099	tggcttggtagacatg	ggaaa-aagaccataatcagggtattg	3139	
Db 42	TGGCGTGTGTTNAG	ATGGAAGNAGACCATCAATCAGGTAGGTG	1	

```

RESULT 2
BM479578      1051 bp      mRNA      linear      EST 05-FEB-2002
LOCUS      AGENCOURT_6464998 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577183
DEFINITION      5', mRNA sequence.
ACCESSION      BM479578
VERSION      BM479578.1 GI:18528620
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1051)
AUTHORS      NTH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12330 row: d column: 16
            High quality sequence stop: 662.

FEATURES             source
     1..1051
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone="IMAGE:5577183"
     /clone_lib="NIH_MGC_92"
     /tissue_type="embryonal"
     /lab_host="DH10B (phage-resistant)"
     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
     Average insert size 2.5 kb. Library enriched for
     full-length clones and constructed by Life Technologies.
     Note: this is a NIH_MGC Library."

BASE COUNT      296 a 278 c 234 g 243 t

ORIGIN
Query Match      23.3%; Score 826.6; DB 10; Length 1051;
Best Local Similarity 95.8%; Pred. No. 1.3e-195;
Matches 892; Conservative 0; Mismatches 34; Indels 5; Gaps 4;

QY 2451 ccaccttaaacacagcttccagcaaacacagcagcaagatgtgccatacaagaact 2510
      |||||||
Db 1 CCACCTTAAACACCAGTTCTCCGACGAAACAGCGCAAGGATGTGCCATACAAAGAACT 60
      |||||||

QY 2511 cttaagcagctgacttccagcagcatcgattctattgacctggggcgaacctttcc 2570
      |||||||
Db 61 CTTAAGCAGCTGACTTCCAGCAGCATGCGATTCTATTGACCTTGGGCGAACCCTTCC 120
      |||||||

QY 2571 tacacaccatactctctgcccagcttgagcagcagcatctcgctttacaacatttt 2630
      |||||||
Db 121 TACACACCATTAATCTCTCTGCCAGCTTGGACGAGGACGATTCGCTTTACAACTTTT 180
      |||||||

QY 2631 gaaggcctactcaactcttagaccaggaagtggatattgccaaaggtctcagctttgtagc 2690
      |||||||
Db 181 GAAGGCTTACTCATTCTAGACCAGGAAGTGGGATATTGCCAAGTCTCAGCTTTGTAGC 240
      |||||||

QY 2691 aggcattttgctcttcattatgagtgaaggagagcgctttaaaaatgctcaagttctcat 2750
      |||||||
Db 241 AGGCATTTTGTCTTCTATATAGTGAGGAAGAGGCGTTTAAATATGCAAGTTTCTCAT 300
      |||||||

QY 2751 gtttaacatgggctcgcgaaacagatcggccagacatgattatttacagatccagat 2810
      |||||||
Db 301 GTTTGACATGGGGCTGCGGAAACAGATATCGGCCACACATGATTATTACAGATCCAGAT 360
      |||||||

```



```
/db_xref="taxon:9606"  
/clone_lib="IMAGE:4851090"  
/tissue_type="primary B-cells from tonsils (cell line)"  
/lab_host="DH10B (phage-resistant)"  
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;  
Site_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH-MGC Library." 171 t  
BASE COUNT 220 a 230 c 214 g 171 t  
ORIGIN  
  
Query Match 20.8%; Score 741.8; DB 10; Length 835;  
Best Local Similarity 98.8%;  
Matches 768; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
  
QY 1591 gatagtccctgtctagatcaatgaacacacagcaagagccatctgtgtgaaag 1650  
Db 1 GATAGCTCCCTGCTAGTACATTAAGTAACACAGCAAGAGCCATCTGTGTGAAAG 60  
QY 1651 gaggccttgccatctctgagagctcttaagctctcggctcctcgaggacctgtcc 1710  
Db 61 GAGGCCCTTGCCCATCTCTGAGAGCTCCTTAAGCTCCTCGCTCCTCGGAGGACCTGTC 120  
QY 1711 agtgactcgagagctatctccagaagagccagctcgctgtgcgccagcaggcttc 1770  
Db 121 AGTGACTCGGAGAGTATCTCCCAAGAGAGCCAGCTCCGTGTGCGCCCGAGGCGCTTC 180  
QY 1771 agggaggcagcaaacacctgagtcacttccctcgaatgacaggaacctccacaacct 1830  
Db 181 AGGAGGCGAGCAACACCTGAGTCACTTCCCATCGAATGCCAGGAACCTCCACACCT 240  
QY 1831 gcccggggtcccggggttgcgaagaaacttatgggtatcactcagtgagca 1890  
Db 241 GCCCGGGGTCCCGGGGGTTCGCAAGGAAACTTTATGAGGTATCACTCAGTGAGACA 300  
QY 1891 gagcgctcatgaacaaaggacttgaatccaaagcaaacctctgtgtattcgt 1950  
Db 301 GAGACGCTCATGAACAAAGGACTTTGAATCCAAAGCAACCATCTGTGTGATCTGGT 360  
QY 1951 gggactctgtgaagaccggagccattcctggaggcagcagatattcctccagtagcc 2010  
Db 361 GGGACTCTGTGAAGACCCGGAGCATCTCTGGAGGCGAGAGATATTCCTCGAGTAGCC 420  
QY 2011 acccgagagcgctgcgattcttcagcagatataagattattcagagctgggagag 2070  
Db 421 ACCCGGAGAGGCGTGCAGTTCTTCAGCAGATATGAAGATTATTACAGAGCTGGGAGAG 480  
QY 2071 ctctcccccagatctctttagaacagcttggagatggcccttggcccccaccca 2130  
Db 481 CTTCCCCCAGATCTCTTTAGAACACAGTTTGTGAAGATGGGCCCTTTGGCCCCCACCAC 540  
QY 2131 gaggaaagaaaaaggacatctcgtgagctcgcagagctgtggcaaaaggctattcttcaa 2190  
Db 541 GAGGAAAGAAAGAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAAAAGGCTATTCTTCAA 600  
QY 2191 cagatactgtcttagaagtggaagaaaaatacagaagctccaagcctctgaaatgat 2250  
Db 601 CAGATACTGTCTGTAGAAATGGAGAGGAAATCAGAAAGCTCCAAAGCTCTGAACATGAT 660  
QY 2251 ttgctgaacagcgctgagctcgattatgaagaaattactcctctgtcttaagaagta 2310  
Db 661 TTGCTGAACAGCGCTGAGCTGCATTATGACGAAATTAATCTCCTGTCTTAAGAAGTA 720  
QY 2311 actacagtgt-gggaaaaagatgctta-gcactccaggaagatcaaaaaattaaagttg 2365  
Db 721 ACTACAGTGTGGGGAACAGATGCTTAGGCACCTCCAGGAGATCAACAATCAGGTTTCG 777
```

RESULT 6
LOCUS BG746377

DEFINITION BG746377 800 bp mRNA linear EST 15-MAY-2001
602703644F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4856891 5',
mRNA sequence.

ACCESSION BG746377
VERSION GI:14057030
KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 800)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-femail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LICM1708 row: h column: 12

High quality sequence stop: 762.

FEATURES

Location/Qualifiers

source

1..800

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="IMAGE:4856891"

/tissue_type="NIH_MGC_15"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)"

BASE COUNT 213 a 218 c 203 g 166 t

ORIGIN

Query Match 20.8%; Score 738.8; DB 10; Length 800;
Best Local Similarity 98.8%;
Matches 787; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

QY 1502 agagatctttaaagagctcttagaagattttgtcccggggtaataaagccagagcc 1561

Db 2 AGAGATCTTTAACAGAGCTTTAGAAAGTATTTGTCCCGGGTAATAAGCCAGAGCC 61

QY 1562 tgcaggaaactccatcagtgatgagctgagctccctgtcttagtacattagtaaca 1621

Db 62 TGCAGGAACACTCCATCAGTGTGATCTGGATAGTCCCTGTCTAGTACATTAGATAACA 121

QY 1622 ccagcaagagccatctgtgtgaaagagggccttgcctctctgagagctccctta 1681

Db 122 CCAGCAAGAGCCATCTGTGTGAAAGAGGCGCTTGCCCATCTCTGAGAGCTCCTTTA 181

QY 1682 agctctcggctcctcgcaggagacctgtccagtgactcggagagtcattctccagaagc 1741

Db 182 AGCTCTCGGCTCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCAGAGAGC 241

QY 1742 cagctcgcgtgtcgccccagcagccttcagggagcgagcaaacacctgagtcactcc 1801

Db 242 CAGCTCGGCTGTGCGCCCGCAGCAGCCCTTCAGGAGCGGCAACACCCCTGAGTCACTTC 301

DEFINITION AGENCOURT_6421241 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531908
5', mRNA sequence.
ACCESSION BM460573
VERSION BM460573.1 GI:18509613
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1109)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: AFCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12214 row: f column: 05
High quality sequence stop: 630.
FEATURES
source
1..1109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5531908"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
BASE COUNT 301 a 311 c 258 g 239 t
ORIGIN
1
|||||
Query Match 20.5%; Score 729.8; DB 10; Length 1109;
Best Local Similarity 97.5%; Pred. No. 2.2e-171;
Matches 773; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
Qy 1340 tgattattttcttgagatgtttatgaagagaaacagaaacacacatccatttg 1399
Db 13 TGATTATTCTTCTGNGATGTTTATGAAGAGAAACAGAAAGACATCCATATTG 72
Qy 1400 gggagatgaagcagacatcgagatggcagcagagaaatttggaagtgaattaccaccca 1459
Db 73 GGGAGATGAAGCAGACATCGCAGATGCGCAGCAGAGAAATATTGGAAGTGAATTACCAACCA 132
Qy 1460 gtgacctgattaggctagatgctgaaacaaacagaaagagatctttaacagagt 1519
Db 133 GTGCCACTCGATTGAGCTAGATGCTGTAAGAAACAAAGCAAGAGATCTTTAACAGAGT 192
Qy 1520 ctttagaaagtattttgccggggtaataaagccagagcctgcaggaacactccatca 1579
Db 193 CTTTAGAAAGTATTGTTCCCGGGTAATAAGCCAGAGCCTGCGAGAACATCCATCA 252
Qy 1580 gtgtggtatcgttagtgcctgtctagtacattaagtaacacagcaagagccatctg 1639
Db 253 GTGTGGATCTGGATAGTCTGCTGTAGTACATTAAGTAACACACAGCAAGAGCCATCTG 312
Qy 1640 tgtgtgaaaggaggccttcccatctctgagagctcctttaagctcctcggtcctcgg 1699
Db 313 TGTGTGAAAGAGAGCGCTCCCATCTCTGAGAGCTCTTTTAAGCTCTCGGCTCTCGG 372
Qy 1700 agacctgtccagtgcctgcgaagctcatctccagagagccagctccgctctcgcccc 1759
Db 373 AGGACCTGTCCAGTGACTCGGAGAGTCTCTCCAGAGAGCCAGCTCCGCTGTGCCCC 432
Qy 1760 agcaggccttcaggaggcgaacacacccctgagtcacttcccccatcgatgcaggaac 1819
|||||

Db 433 AGCAGGCCTTCAGAGGCGGAGCAACACCCTGAGTCACCTCCCATCGAATGCCAGAAC 492
Qy 1820 ctccacaacctgccgggggtcccccgggttttcgaaagaaacttatgaggtatccact 1879
Db 493 CTCACAACTGCGCGGGGTTCGCCGGGGTTCGCAAGAGAAACTTATGAGGTATCATC 552
Qy 1880 cagtgcagcagagagcctcatgaacgaagagactttgaatccaaagcaaacactcttg 1939
Db 553 CAGTGACACAGAGACCCCTCATGAACGAAGGACTTTGAATCCAAAGCAACCATCTTG 612
Qy 1940 gtgattctgtggtgactcctgtgaagaccggaggcattcctggagcagcagatattcc 1999
Db 613 GTGATCTGCTGGACTCCTGTGAAGACCCCGGAGCATTCCTGGAGCAGCATATTCC 672
Qy 2000 tcgagtagccaccccgagaaagcgctgcgattcttccacagatagaagattatccag 2059
Db 673 TCCGAGTAGCACCCCGCAGAGCGGGCGATTCTTCACAGATATGGAGATTATTCGG 732
Qy 2060 agct-gggagagcttccccccacgacg-cctcttagaacaggtttgt-gaagatggccct 2116
Db 733 AGCTGGGAGAGCTTCCCCCCCCCAACTTCCCTTAGAACCAAGTTGTGGAAGATGGGGCC 792
Qy 2117 ttggccccccacc 2129
Db 793 TTTGGCGCCTCC 805
RESULT 9
BE299948
LOCUS 600944494F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960782 5',
DEFINITION mRNA sequence.
ACCESSION BE299948
VERSION BE299948.1 GI:9183696
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 952)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: AFCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM54 row: k column: 23
High quality sequence stop: 709.
FEATURES
Location/Qualifiers
1..952
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2960782"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; CDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 249 a 240 c 250 g 212 t 1 others
ORIGIN

Query Match 20.3%; Score 722.2; DB 10; Length 952;
Best Local Similarity 97.2%; Pred. No. 1.7e-169;
Matches 766; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

QY 1879 tcaagtacacagacgacctatgaacaaagaggaactttgaatccaaagcaaacattctt 1938
DB 1 TCAGTGAGCAGACAGAGCGCTCATGAACAAAGAGACTTTGAATCCAAACCAACCATCTT 60

QY 1939 ggtgattctgggtgggactcctgtgaagaccgagagcattctctggagcgagcagatattc 1998
DB 61 GGTGATTCTGGTGGGACCTCTGTGAAGACCCGGAGGCATCTCTGGAGCGAGCAGATATTC 120

QY 1999 ctccgagtagccaccgcccgagagggcgtgctgattcttccagcagatatgaagattattca 2058
DB 121 CTCGGAGTAGCCACCACCCGAGAGGGGTGCGATTCTTCAGCAGAGATATGAAGATTATTC 180

QY 2059 gagctgggagagcttccccccagatctcttttagaaccagtttgaaagatgggcccctt 2118
DB 181 GAGCTGGGAGAGCTTCCCCCAGCATCTCTTTAGAACCAAGTTGTGAAGATGGGCCCTTT 240

QY 2119 ggcctcccccacagagaaagaaagacatctctgtgagctccgagagctgtggcaaaag 2178
DB 241 GGGCCCCCACCAGAGAAAGAAAGAGACATCTCTGAGCTCCGAGAGCTGTGGCAAAAG 300

QY 2179 gctattcttcaacagatactgctgcttagaattggagaagaaatcagaagctccaagcc 2238
DB 301 GCTATTCTTCAACAGATAGCTGCTGTAGATGGAGAGAAATCAGAAGCTTCCAAGCC 360

QY 2239 tctgaaatgatttctgaacagcgcctgaagctcgattatgaagaataattcctcctgt 2298
DB 361 TCTGAAATATGATTGCTGAACAAAGCCCTGAAGCTCGATTATGAAGAAATTAATCCTCCTGT 420

QY 2299 cttaagaagtaactacagctgtgggaaagatgcttagcactccaggaagatcaaaaatt 2358
DB 421 CTTAAGAGAGTACTACAGTGTGGGAAGAGATGCTTTAGCACTCCAGGAAGATCAAAAT 480

QY 2359 aggtttgacatggaaaaatgcaactgctgtgttgggcaagggtgtgccagctcaccga 2418
DB 481 AAGTTTGACATGGAAAAATGCACTCGGCTG-TGGCAAGGTGTGCCAGCTATCACCGA 539

QY 2419 ggtgaacttgaaattctagctgagcaattccaccttaaacaccagtttcccagcaa 2478
DB 540 GGTGAATCTGGAAATTTCTAGCTGAGCAATTCACCTTTAAACACCAGTTTCCAGCAA 599

QY 2479 cagcagcaaaagatgtgccatacaagaactcttaaaagcagctgacttcccagcagcat 2538
DB 600 CAGCAGCCAAAGGATGTGCCATACAAAGAACTTTAAAGCAGCTGATCTCCAGCAGCAT 659

QY 2539 ggaattctta-ttgacctggcggaacaccttctctacacaccatacttctctgcccagct 2597
DB 660 GCGATTCTTATTGTGACCTGNGCGAACCTTTCTACACACCACCATACTTCTCTGCCAG-T 718

QY 2598 tgagcagacagctatcgtcttacaacattttgaaggcctactcactcttagaccagga 2657
DB 719 TGAGCAGACAGACATATCGCTTTACAAACATTTGTGAGGGCTACTCCACTCTTAGACAGGA 778

QY 2658 agtgggat 2665
DB 779 GTGGGTAT 786

RESULT 10
BI907513
LOCUS 603065640F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214945 5',
DEFINITION mRNA sequence.
ACCESSION BI907513
VERSION BI907513.1 GI:16170347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 838)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9pbbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11539 row: o column: 10
High quality sequence start: 4
High quality sequence stop: 780.

FEATURES
source
1..838
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5214945"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: PCWV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

BASE COUNT 237 a 199 c 191 g 211 t
ORIGIN

Query Match 20.1%; Score 713.8; DB 10; Length 838;
Best Local Similarity 97.0%; Pred. No. 2e-167;
Matches 812; Conservative 0; Mismatches 17; Indels 8; Gaps 8;

QY 2226 gaagctcaagcctctgaaatgatttctgaacagccctgaagctcgattatgaaga 2285
DB 9 GAAGCTCCAAGCCTCTGAAATGATTGCTGAACAAAGCCCTGAAGCTCGATTATGAAGA 68

QY 2286 aattactcctctcttaagaagtaactacagctgtgggaaagatgcttagcactccagg 2345
DB 69 AATTACTCCTCTCTTAAGAAGTAACTACAGTGTGGGAAAGATGCTTAGCAGCTCCAGG 128

QY 2346 aagatcaaaaaataagtttgacatggaaaaaatgcaactcggtgttgggcaagggtgccc 2405
DB 129 AAGATCAAAAATTAAGTTTGACATGGAAAAATGCACTCGGCTGTGGCGAAGGTGTGCC 188

QY 2406 acgtcatcacagctgaaatctgaaatttctagctgagcaattccaccttaaacacca 2465
DB 189 AGCTCATCACCGAGGTGAATCTGGAAATTTCTAGCTGAGCAATTCACCTTTAAACACCA 248

QY 2466 gttccccagcaaacagcagccaaaggatgtgccatacaagaactcttaaacagctgac 2525
DB 249 GTTTCCACGACAAACAGCAGCCAAAGGATGTGCATACAAAGAACTCTTAAAGCAGCTGAC 308

QY 2526 ttccccagcagctgagattcttattgaccttggggaaaccttctcacacccattct 2585
DB 309 TTCCCAGCAGCATGGCGATTCTTA-TGACCTTGGCGAAACCTTTCCCTACACACCATCT 367

QY 2586 cctctgcccagcttgagcagcagcagctatcgctttacaacattttgaagccttactcact 2645
DB 368 CTCTGCCAGCTTGGAGCAGGACAGCTATCGCTTTTACAACA-TTTGAAGCCCTTACTCACT 426

QY 2646 tctagaccaggaagtgaggatattgccaaggtctcagctttgtagcagggatttggctct 2705
DB 427 TCTAGACCAAGGAATGGGATA-TGCCAAGGTCTCAGC-TTGTAGCAGGCA-TTTGTCTCT 483

Qy 2706 tcataatgaggaagcgggttttaaaatgctcaagtttctgtatgtttgacatgggct 2765
 Db 484 TCATATGAGTGAGGAGAGGCGTTAAATGCTCAAGTTCTGATG-TTGACATGGGGCT 542
 Qy 2766 gcgaaacagtcagcggcagacatgattttacagatccagatgtaccagctctcag 2825
 Db 543 GCGAAACAGTATCGGCAGACATGATTTATTTACAGATCCAGATGTACCAGCTCTCGAG 602
 Qy 2826 gtcttcacatgattacacagagaccctcacaatcaactcgagagacacagatcgccc 2885
 Db 603 GTTGCTTCATGATTACACAGAGACCTCTACAATCACCTGGAGAGACAGATCGGCC 662
 Qy 2886 cagctctacgctg-cccccctggctctcaccatgttgcctcaccagttcccgctggat 2944
 Db 663 CAGCTCTACGCTCGCCCTGGTCTCACCAGTTTGGCTCACATTCGCCGTGGAT 722
 Qy 2945 tcgtagccagagtccttgatagattttcttcaggggaacagaggtcatattaaagt 3004
 Db 723 TCGTAGCCAGAGCTTTTGATGATGATTTTCTTCAGGGAACAGAGGTCATATTAAAGTGG 782
 Qy 3005 ctttaagctgttgggaagccataagcccttgattctgcagcatgaaacccatgaaa 3061
 Db 783 -TTTAAATCTGTGGGAGGCGATAAGCCCTGATTCTGCAGCATTTAAACCTAGAAA 838

RESULT 11
 BG120498 807 bp mRNA linear EST 30-JAN-2001
 LOCUS 602346815F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4441923 5',
 DEFINITION mRNA sequence.
 ACCESSION BG120498
 VERSION BG120498.1 GI:12614007
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 807)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.lnl.gov>
 Plate: LLAM10213 row: n column: 04
 High quality sequence stop: 682.
 Location/Qualifiers
 1..807
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4441923"
 /clone_lib="NIH_MGC_90"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 231 a 197 c 189 g 190 t
 ORIGIN

Query Match 19.8%; Score 705; DB 10; Length 807;
 Best Local Similarity 96.1%; Pred. No. 3.le-165;
 Matches 766; Conservative 0; Mismatches 25; Indels 6; Gaps 4;

Qy 1992 gatattctccagtagccaccocgcagaaagcgtgcgattcttccagcagatatgaaga 2051
 Db 1 GATATTCTCCGAGTAGCCACCCCGCAGAGGCGTGGATTTCTCCAGCAGATATGAGA 60
 Qy 2052 ttattcagagctggagagcttccccccacgatctcctttagaaccagttgtgaagtgg 2111
 Db 61 TTATTACAGAGCTGGAGAGCTTCCCCCAGCATCTCCTTTAGAACCACTTTGTGAAGTGG 120
 Qy 2112 gccctttggccccccacagagagaaagagacatctcgtgagctcagagagctgtg 2171
 Db 121 GCCCTTTGGCCCCCACCAGAGGAGAAAGAGACATCTCGTGAGCTCCGAGAGCTGTG 180
 Qy 2172 gcaaaagcgtattctcaacagatcactgctgttagaatgagaagaaatcagaagct 2231
 Db 181 GCAAAAGCCTATTCTTCACAGATACCTGCTTGAGTAATGGAGAGAAATACAGAGCT 240
 Qy 2232 ccaagcctctgaaatgatttgcctgaacagcgcctgaagctcgattatgaagaaatc 2291
 Db 241 CCAAGCCTCTGAAATGATTGCTGGAACAAGCGCTCAAGCTCGATTATGAAGAAATAC 300
 Qy 2292 tcctgtcttaagaagtaactacagtggtgggaaagatccttagcactccaggaagatc 2351
 Db 301 TCCCTGTCTTAAAGAGTAATACTACAGTGTGGGAAAGATGCTTAGCACTCCAGGAAGATC 360
 Qy 2352 aaaaattaaatttgacatggaaaaatgcactcgctgttgggcaaggtgtgccacgtca 2411
 Db 361 AAAAAATTAAAGTTGACATGGAATAATGCACCTCGCTGTTGGCAAGGTGTGCCACGTCA 420
 Qy 2412 tcaccgaggtgaaatctggaattcttagctgagcaattccaccttaaacacacagttcc 2471
 Db 421 TCACCGAGGTGAAATCTGAAATTTCTAGCTGAGCAATTCACCTTAAACACACAGTTTCC 480
 Qy 2472 cagcaaacagcagccaaagatgtgccatacaagaagactttaagcagctgacctccca 2531
 Db 481 CAGCAAAACAGCAGCAAAAGGATGTCCTACAAAGAACTCTTAAAGCAGCTGACTTCCCA 540
 Qy 2532 gcagctgcgattcttattgaccttggcgaaacctttccacaca-cccatacttctctg 2590
 Db 541 GCAGCATGCGATCTTATTGACCTTGGCGAAGCTTTCTTACACAGCCCATACTTCTCTG 600
 Qy 2591 ccagcttgagcagcagcagctatcgctttacaacatttggaagcctactactctctag 2650
 Db 601 CCCAGCTTGGAGCAGGACAGCTATCGCTTTTACAACTTTTGAAGGCTTACTACTTCTAG 660
 Qy 2651 accaggaagtggaatattgcc--aaggtcagcttt-gtagcagcatittgtctcttc 2707
 Db 661 ACCAGGAAGTGGGATATGCCCCAAAGTCTTCAAGCTTTGGTGGCAGGCTCTTGTCTCTC 720
 Qy 2708 atatgagtgaggaagcggctttaaagtctcaagtttctgattttgacatggggctgc 2767
 Db 721 ATATGAGTGACGAGAGGCGTTAAATGCTCCA--GTTCCGATGTTTGACATGGGCCGCG 778
 Qy 2768 ggaacagtatcggcca 2784
 Db 779 GGAACCAAGTTCCGGCAA 795

RESULT 12
 BM014940 832 bp mRNA linear EST 30-OCT-2001
 LOCUS 603640877F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5416818 5',
 DEFINITION mRNA sequence.
 ACCESSION BM014940
 VERSION BM014940.1 GI:16529294
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 832)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/Dnp
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12063 row: j column: 19
 High quality sequence stop: 811.
 Location/Qualifiers
 1. 832
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5416818"
 /clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="organ: breast; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC library."
 231 a 205 c 186 g 208 t 2 others

FEATURES
source

Query Match 19.7%; Score 701; DB 10; Length 832;

Best Local Similarity 97.5%; Pred. No. 3.1e-164; Mismatches 0; Indels 8; Gaps 7;

QY 2405 cagctcatcaccgagtgaaatctggaattctagctgagcaattccacccttaaacacc 2464

DB 29 CAGGTATCACCAGGTGAATCTGGAAATTTCTAGCTGAGCAATCCACCTTAACACC 88

QY 2465 agtttcccagcaaacagcagcgcgaaggatgtgccatacaaaagaactcttaagcagctga 2524

DB 89 AGTTTCCCAGCAACAGCAGCAAGGATGTGCCATACAAAGAACTTTAAAGCAGCTGA 148

QY 2525 ctcccagcagcatcgattcttattgacctggcggaacctttctcacaccatact 2584

DB 149 CTCTCCAGCAGCATGGATCTCTTA-TGACCTTGGGGAACCTTTCTTACACACCCATACT 207

QY 2585 tctctgccagcttgagcagcagcagctatcgcttttaacaacattttgaaggcctactcac 2644

DB 208 TCTCTGCCAGCTTGGAGCAGCAGCAGCTATCGCTTTTACAACTTTTGAAGGCTACTCAC 267

QY 2645 ttctagaccaggaagtgggatatgtcccaaggtctcagcttttagcaggcatttgcctc 2704

DB 268 TTCTAGACCAGGAAGTGGGATATTGCCAAGGCTCAGCTTTGTAGCAGGCATTNTGCTTC 327

QY 2705 ttcatatgagtgaagaaggcggtttaaaatgctcaagtttctgatgtttgacatggggc 2764

DB 328 TTCTATGATGAGGAGGAGGCGTTTAAATGCTCAAGTTTCTGATGTTGACATGGGGC 387

QY 2765 tgcgaaacagctatcgccgagacatgattattttacagatccagatgtaccagctctga 2824

DB 388 TCGGAAACAGTATCGCCAGACATGATTTTACAGATCCAGATGTACCAGCTCTCGA 447

QY 2825 gtttgcctcatgattaccacagagaccttacaataccctggaggagcagcagatcgccc 2884

DB 448 GGTGTGCTTATGATTACCAACAGAGACCTCTACAATCACCTGGAGGAGCAGATCGGCC 507

QY 2885 ccagctctacgctcccctgttctcaccatgtttgcttccacagttcccgctgggat 2944

DB 508 CCAGCCTCTACGCTGCCCTGGTTCCTCACCATGTTTCCTCAGATTTCCCGCTGGGAT 567

QY 2945 tcgtagccagagcttttgatagatttttcttcagggaacagaggtcatatttaa-agtg 3003

DB 568 TCGTAGCCAGAGCTTTGATATGATTTTCTTCAGGGAACAGAGTTCATATTTAACAGTG 627

QY 3004 gctttaagtctgttgggaagccataagcccttgattctgcagcagtaaaacccatagaacc 3063

DB 628 GCTTTAAGTCTGNTGGGAAGCCATAAGCCCTTGATTTCTGCAGCATGAACACCTAGAAACC 687

QY 3064 atagtgtacttat-aaaaacacgctaccacaccttggc-ttggtacagat-ggaaaag 3120

DB 688 ATAGTGTACTTTATATAAAAGCAGCGCTAACCAAGCTTGCTTTGTACAGATGGGAAAG 747

QY 3121 accatcaatcaggtatttgaaatggacatcgctaaa--cagttacaagcttatgaagtgg 3178

DB 748 ACCATCAATCAGGTATTGGAATGGACATCGGTAAACAGTTACAAAGCTTATGAAGTTG 807

QY 3179 agtacc-acgtctctcaagaagaac 3202

DB 808 AGTACCAACGTCCTCTCAAGAAGAC 832

RESULT 13

AL570425 851 bp mRNA linear EST 16-FEB-2001
 LOCUS AL570425 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1020YF02 5
 DEFINITION prime, mRNA sequence.

ACCESSION AL570425
 VERSION AL570425.1 GI:12926720

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 851)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 Location/Qualifiers

1. 851

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1020YF02"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 240 a 214 c 220 g 174 t 3 others

ORIGIN

Query Match 19.4%; Score 689.8; DB 9; Length 851;

Best Local Similarity 98.9%; Pred. No. 2e-161;

Mismatches 1; Indels 3; Gaps 2;

Matches 714; Conservative 1; Mismatches 4; Indels 3; Gaps 2;

QY 2836 gattaccacagagacctctacaat--cacctggagagcagcagatcgccccagcctct 2893

DB 1 GATTACCACAGAGACTCTACAANTCACCTTGGAGGAGCAGAGATCGGCCCGAGCTCT 60

QY 2894 acgtgccccctgttctcaccatgtttgctccagttccgctgggattcgtagcca 2953

DB 61 AGCGTGCCCCCTGGTCTCCACCATGTTTGCCTCACAGTCCCGCTGGGATTCGTAGCCA 120

QY 2954 gagtcttgatatgattttcttcagggaacagaggtcatatttaaaagtggcttaagtc 3013

DB 121 GAGTCTTGATATGATTTTCTTCAGGGAACAGAGTCATATTTAAAGTGGCTTTAAGTC 180

```
QY 3014 ttttgggaagccataagcccttgattctgcagcatgaaacccctagaaacccatagttgact 3073
D 181 TGTGGGAAGCAYTAAGCCCTTGATTCGACAGTGAAGAACTAGAAACCAATAGTTGACT 240
QY 3074 ttataaaagcagcgtactaccacccctggcttggtacagatgaaagacccatcaatcagg 3133
D 241 TTATAAAAGCAGCGTACCCCAACCTTGCTTGGTACAGATGGAAGAACCAATCAATCAGG 300
QY 3134 ttttgaatggaatgcgtcaaaacagttacaagcttatgaagcttgatgaagctacacgtcttc 3193
D 301 TATTGGAATGGATCGCTAATACAGTTTACAGCTTATGAAGTTGAGTACCACTGCTCCCTC 360
QY 3194 aagaagaacttatgattctctctctcctcagtgagacacaaagaatgataaattagaga 3253
D 361 AAGAAGAATTTATGATTCCTCTCTCTCAGTGACACCAAGAAATGGATAATTAGAGA 420
QY 3254 aaacacacagcagcttacgaaacagaacaccttgacctcttgaaacagttgcagtgagcaa 3313
D 421 AAACCAACAGCAGCTTACGAAACAGAACCTTGACCTCTTGAACAGTTGCAAGGTGGCAA 480
QY 3314 atggtagatccaagccttgagccaccattgagaagctcctgagcagtgagagcaagc 3373
D 481 ATGGTAGATCCAAAGCCTTGAGCCACCATTTGAGAGCTCCTTGAGCAGTGAGAGCAAGC 540
QY 3374 tgaagcagcgcatgcttaccttagaactggagcgttgctggcctctgctgcagacggtgagg 3433
D 541 TGAAGCAGCGCATGCTTACCTTAGAAGTGGAGCGGTGGCCCTGCTGCAGACGTTGAGG 600
QY 3434 agctggcggcggagcagcagcccgagcggcggcggcggcggcggcggcggcggcggcggc 3493
D 601 AGCTGGCGCGCGGA-SGCAGAGCCAGCGAGCGGAGCTGAGTGACGCGAGCGCGGAGC 659
QY 3494 ccagggcgactgacagctctgcaggagagattgcaacacacacacacacacacacacacac 3553
D 660 CCAGGGCGAGTACAGCTCTCAGGAGAGATTGCAACACCATCCACACTGTCAGGCC 719
QY 3554 tt 3555
D 720 TT 721

RESULT 14
BE300005 60944594F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960783 5',
LOCUS mRNA linear EST 20-JUL-2000
DEFINITION mRNA sequence.
ACCESSION BE300005
VERSION BE300005.1 GI:9183753
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cygabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: L1CM54 row: k column: 24
High quality sequence start: 2
High quality sequence stop: 753.
Location/Qualifiers
1. 754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2960783"
```

```
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pORF7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT	227 a	186 c	177 g	164 t
ORIGIN				

Query Match	19.2%	Score	683.4	DB	10	Length	754		
Best Local Similarity	99.1%	Pred. No.	7.5e-160						
Matches	750	Conservative	0	Mismatches	1	Indels	6	Gaps	6

QY	1878	ctcagtcagcacagacgacctcatgaacgaagacattgaatccaaagcaaacatctc	1937
D	1	CTCAGTGAGCACAGAGAGCGCTCATGAACGAAGGACTTTGAATCCAAAGCAACCATC	59
QY	1938	tgggtattctggtggactcctgtgaagaccggaggcattctctggaggcagcatatt	1997
D	60	TGGTGATTCGTGGGACTCCTGTGAAGACCCGGAGGCAATTCCTGGAGGCAGAGATAT	119
QY	1998	cctccgagtagccaccgccgagaggcgtgcgattcttccagcagatatgaagtattc	2057
D	120	CCTCCGAGTAGCCACCCCGCAGAGGCGTGGATTCTTCCAGCAGATATGAAGATTATTC	179
QY	2058	agaactgggagagcttccccacgactctctttagaacagcttggtaagatggccctt	2117
D	180	AGAGCTGGGAGAGCTTCCCCACGATCTCCTTTAGAACCCAG-TTGTGAAGATGGGCC	237
QY	2118	tggcccccaccagaggaagaaagagacatctctgagctccgagagctgtggcaaaa	2177
D	238	TGGCCCCCACCAGAGGAAAGAAAGACATCTCGTGAGCTCCGAGAGCTGTGCAAAA	297
QY	2178	ggctattcttcaacagatactgctgcttagaagtggaaggaagaaatcagaagctccaa	2237
D	298	GGCTATTCTTCAACAGATACTGCTGCTTAGAATGGAAGGAAATCAGAAGCTCCAAGC	357
QY	2238	ctctgaaatgattctgaacagccctggaagctcgatttatgaagaaatctactcctg	2297
D	358	CTCTGAAAAATGA-TTGCTGAACAAGCGCTGAAAGCTCGATTATGAGAAATTTACTCCCTG	416
QY	2298	tcttaagaagtaactacagtggtgggaaagatgcttagcactccaggaagatcaaaaa	2357
D	417	TCTTAAAGAAAGTAATACTACAGTGTGGGAAAGATGCTTAGCAGCTCCAGAGATCAAAAAT	476
QY	2358	taagtttgacatggaaaaaatgcaactggctgttgggcaaggtgtgcacgtcatcccg	2417
D	477	TAAAGTTGACATGGAAAAAATGCACTCGGTGTGGCAAGGTGTGCCAGCTCATACCCG	536
QY	2418	aggTgaatctggaaatttctagctgagcaattccacttaaacaccagtttccacagcaa	2477
D	537	AGGTGAATCTGGAAATTTCTAGCTGAGCAATTCACCTTAAACACCAGTTTCCACGAA	596
QY	2478	acagcagcgaagagatgtgcatacacaagaactcttaagaagctgacttcccagcagca	2537
D	597	ACAGCAGCAAAAGGATGTGCATACAAAGAACTCTTTAAAGCAGCTGACTTCCCAAGCAGCA	656
QY	2538	tggattcttatgaccttggggaacctttcttaacacccataacttctctgcccagct	2597
D	657	TGGGATTTCTTA-TGACCTTGGGGAACTTCTCTTACACACCCACTACTTCTTGTGCCAGCT	715
QY	2598	tggagcaggagcagctatcgctttacaacatttgaag	2634
D	716	TGGAGCAGGACAGCTATCGC-TTACAACATTTTGAGG	751

RESULT 15

BE561889

LOCUS 601346094F1 NIH_MGC_8 892 bp mRNA linear EST 15-AUG-2000
DEFINITION 601346094F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3678887 5',
mRNA sequence.

ACCESSION

BE561889

VERSION

BE561889.1 GI:9805609

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 892)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D.

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LCM356 row: 11 column: 24

High quality sequence start: 50

High quality sequence stop: 758.

Location/Qualifiers

1..892

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3678887"

/clone.lib="NIH_MGC_8"

/tissue.type="Burkitt lymphoma"

/lab.host="DH10B (phage-resistant)"

/note="Organ: lymph; Vector: pORF7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GCGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 224 a 240 c 242 g 186 t

ORIGIN

Query Match 19.2%; Score 682.6; DB 10; Length 892;

Best Local Similarity 95.3%; Pred. No. 1.3e-159;

Matches 790; Conservative 0; Mismatches 29; Indels 10; Gaps 8;

Qy 1628 aagagccatctgtgtgaaaggagccttgcacatctctgagagctcctttaagctcc 1687

Db 63 AAGAGCCATCTGTGTGAAAGGAGG-CTTCCCATCTCTGAGAGCTCTTTAAGCTCC 121

Qy 1688 tcggctccctcgagagacctgtcagtgactcgagagagtcacccagagagccagctc 1747

Db 122 TCGGCTCCTCGGAGGACCTGTCCAGTACTCGGAGAGTATCTCCAGAGAGCCAGCTC 181

Qy 1748 cgctgtcgccccagcaggccttcaggaggcgagcaaacacccttgagtcactt-ccccatc 1806

Db 182 CGCTGTGCGCCCGCAGAGGCTTCAGGAGCGAGCAACACCCCTGAGTCACTTCCGCCCATC 241

Qy 1807 gaatgccagaacctccacaacctcccggggtcccggggttttcgcaaggaacctt 1866

Db 242 GAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCCGGGGGTTCGCAAGGAACCTT 301

Qy 1867 atgaggtatcactcagtgagcagagacgctcatgacgaagagactttgaatccaaa 1926

Db 302 ATGAGGTATCATCTAGTGACAGAGAGCGCTCATGACGAAGAGGACTTTGATCCAAA 361

Qy 1927 gcaaacattctgtgtgattctgtgtggactcctgtgaagaccgccgagggattcctcggagg 1986

Db 362 GCAAAATTATCTTGGTGATTCTGTGGGACTCTCTGTGAAGACCCGGAGGCATCTCTGGAGG 421

Qy 1987 cagcagatatctctccagtagccaccaccccgagagggcgctgcgattcttccagcagatat 2046
Db 422 CAGCAGATATTCTCTCCGAGTAGCCACCCTCCGAGAGGGGTGGATTTCTTCCAGCAGATAT 481
Qy 2047 gaagattattcagagctgggagagcttcccccaagatctctcttagaaccagtttggtaa 2106
Db 482 GAAGATTATTAGAGCTGGGAGAGCTTCCCGCAGCATCTCCTTTAGAACCAAGTTTGTGAA 541
Qy 2107 gatgggccccttggccccccacccagagagaaagagacatctcgtgagctcggagag 2166
Db 542 GATGGGCCCTTTGGCCCCCACCAGAGAGAAAGAGATCTCGTGAGCTCCGAGAG 601
Qy 2167 ctgtggcaaaaggctattcttcaacagatactgctgcttagaagagaaagaaatcag 2226
Db 602 CTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGCTTAGAATGCGCAAGGAAATCAG 661
Qy 2227 aagctcccaagcctctgaaaaatgatttctgaacaaagcgcctgaagctcgattatgaagaa 2286
Db 662 AAGCTCCAAGCCTCTGAAAATGATTTGCTGAACAAGCGCCTGAAGCTCGATTATGAAGAA 721
Qy 2287 attactccctg-tcttaagaagtaactacagttgtgga--aaagatgcttagcactcca 2343
Db 722 ATTACTCCCTGTCTTAAAGAAAGTAACACAGTGTGGGACCAAGATGGCTTAGCACTCCA 781
Qy 2344 ggaagatcaaaaaattaaagtcttgcacatggaaaaaatgcactcggtgttgggcaagtggtg 2403
Db 782 GGAAGATC-CCCATTTAGTTTGGCCTGCTGACAATGC-CTCGGGGTGTTGGGCCAGGGGG 839
Qy 2404 ccagctcatcccgaggtgaaatctcgaaaattcttagctgagcaattcc 2452
Db 840 CC-CGTGTCATCCCGCGCTGAAACTGGG--AGTTTTAGTGGGCAATTC 885

Search completed: September 9, 2002, 16:41:28
Job time: 9256 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: September 9, 2002, 19:18:41 : Search time 706.85 Seconds
(without alignments)
8634.973 Million cell updates/sec

Title: US-09-762-311-4_COPY_176_3730

Perfect score: 3555

Sequence: 1 atggaaccaataacattcac.....tccacactgtccagccctt 3555

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/hold-geneseq/geneqseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3555	100.0	3983	21	AAZ50906 Human TBC-1 cDNA f
2	3555	100.0	3988	21	AAZ50907 Human TBC-1 cDNA f
3	3026.4	85.1	3326	22	AAI57854 Human polynucleoti
4	2940.2	82.7	3727	21	AACT7404 Human OREF2959
5	2833.8	79.7	3281	22	AAI59640 Human polynucleoti
6	2510.4	70.6	4039	19	AAV05886 Human tbc-1 gene.
7	2112	59.4	2362	22	AAH18401 Human cDNA sequenc
8	1000.8	28.2	1422	22	AAI34787 cDNA encoding nove
9	999.4	28.1	2989	22	AAK94796 Human full-length

10	943.2	26.5	1269	22	AAI34785	cDNA encoding nove
11	785	22.1	883	22	AAH04450	Human CDNA clone (
12	574.2	16.2	849	22	AAK92006	Human CDNA 5'-end
13	574.2	16.2	849	22	AAK93666	Human TBC-1 partia
14	422.6	11.9	17590	21	AAZ50904	Human gene express
15	362	10.2	696	20	AAI15229	Drosophila melanog
16	306.2	8.6	5472	23	ABL02013	Human TBC-1 partia
17	284.4	8.0	99960	21	AAZ50905	Human secreted pro
18	280	7.9	458	20	AAI40323	Drosophila melanog
19	245.6	6.9	8105	23	ABL02012	Human CDNA clone (
20	228.4	6.4	555	22	AAH09958	Human RECAP polynu
21	132.8	3.7	3691	22	AAI59516	Human polynucleoti
22	132.8	3.7	4827	22	AAI59592	Human polynucleoti
23	132.8	3.7	5294	22	AAI57806	Drosophila melanog
24	128.8	3.6	3585	23	ABL12337	CDNA #1312 encodin
25	123.8	3.5	154	23	AAI58636	Drosophila melanog
26	119.4	3.4	2626	23	ABL18133	Drosophila melanog
27	119.4	3.4	7829	23	ABL18132	Drosophila melanog
28	108	3.0	2424	23	AAI71793	Arabidopsis thalia
29	83.8	2.4	1601	21	AAI33401	Drosophila melanog
30	79.6	2.2	6765	23	ABL12336	Human gene express
31	78.4	2.2	300	20	AAI13067	Arabidopsis thalia
32	75.2	2.1	1452	21	AAI38750	Human secreted pro
33	73.8	2.1	1140	21	AAI97020	Human secreted pro
34	68.4	1.9	3342	22	AAI32789	Human cancer assoc
35	68	1.9	2492	21	AAI78199	Human secreted pro
36	65.2	1.8	1831	22	AAI64964	Human polynucleoti
37	65	1.8	1117	22	AAI61302	Human cell cycle a
38	65	1.8	2035	22	AAI59600	Human polynucleoti
39	65	1.8	2482	22	AAI59516	Human CDNA sequenc
40	65	1.8	2666	22	AAH17866	Drosophila melanog
41	62.6	1.8	2531	23	ABL07359	Drosophila melanog
42	62.6	1.8	4661	22	ABL07358	Human CDNA encodin
43	61.6	1.7	2801	22	AAI32643	CDNA encoding nove
44	61.6	1.7	2803	22	AAI27080	Human polynucleoti
45	61.6	1.7	2903	22	AAI51641	

ALIGNMENTS

RESULT 1

AAZ50906

ID AAZ50906 standard; cDNA; 3983 BP.

XX AAZ50906;

AC AAZ50906;

XX 31-MAY-2000 (first entry)

XX Human TBC-1 cDNA from first transcript.

XX TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;

XX Single nucleotide polymorphism; tissue differentiation; prostate cancer;

XX Linkage analysis; genetic map; detection; diagnosis; genotyping;

XX transgenic animal; screening; alternative splicing; ss.

XX Homo sapiens.

XX Key

XX 5'UTR Location/Qualifiers

FT 1..170

FT /*tag= a

FT /note= "Spans through exon 1 and part of exon 2"

FT 171..3677

FT /*tag= b

FT /product= "Human TBC-1 protein"

FT /note= "First transcription product"

FT 3736..3983

FT /*tag= c

FT polyA_signal

FT 3942..3947

FT /*tag= d

XX WO200008209-A2.

Db 1731 ctgcaggaaactccatcagctggatctgtagtccctctgcttagtacctaaagtaac 1790
Qy 1621 accagcaagagccatctgtctgtgaaaaaggagcccttgcctatctctgagagctccttt 1680
Db 1791 accagcaagagccatctgtctgtgaaaaaggagcccttgcctatctctgagagctccttt 1850
Qy 1681 aagctccctcgctctcggaggaacctgtccagtaactcggagagatcatctcccaagag 1740
Db 1851 aagctccctcgctctcggaggaacctgtccagtaactcggagagatcatctcccaagag 1910
Qy 1741 ccagctccgctgtgcgccagcaggccttcaggagggcgagcaaacacccctgagtcacttc 1800
Db 1911 ccagctccgctgtgcgccagcaggccttcaggagggcgagcaaacacccctgagtcacttc 1970
Qy 1801 cccatcgaaatccaggaacctccacaacctcccgccgggggtcccggggtttcccaagg 1860
Db 1971 cccatcgaaatccaggaacctccacaacctcccgccgggggtcccggggtttcccaagg 2030
Qy 1861 aaacttaagaggtatcactcagtgagcacagagacgcctcatgaacgaagagactttgaa 1920
Db 2031 aaacttaagaggtatcactcagtgagcacagagacgcctcatgaacgaagagactttgaa 2090
Qy 1921 tccaaagcaaacactcttggtgattctggtgggactcctgtgaagacccggagcattcc 1980
Db 2091 tccaaagcaaacactcttggtgattctggtgggactcctgtgaagacccggagcattcc 2150
Qy 1981 tggagcgacagatattctcctcagtagtaccaccccgagagggcggtgcgattcttcacg 2040
Db 2151 tggagcgacagatattctcctcagtagtaccaccccgagagggcggtgcgattcttcacg 2210
Qy 2041 agatatgaagattatcagagctggagagcttcccacagatctcctttagaaccagtt 2100
Db 2211 agaatagaagattatcagagctggagagcttcccacagatctcctttagaaccagtt 2270
Qy 2101 tgtgaagatggcccttggcccccacagagagaaaaagagacatctcgtgagctc 2160
Db 2271 tgtgaagatggcccttggcccccacagagagaaaaagagacatctcgtgagctc 2330
Qy 2161 cgagagctgtggcaaaaagctattcttcaacagatctcgtcttagaattggagaagaa 2220
Db 2331 cgagagctgtggcaaaaagctattcttcaacagatctcgtcttagaattggagaagaa 2390
Qy 2221 aatcagaagctccagcctctgaaaatgatttctgcaacagcgcctgaaagctcgattat 2280
Db 2391 aatcagaagctccagcctctgaaaatgatttctgcaacagcgcctgaaagctcgattat 2450
Qy 2281 gaagaaattactcctctcttaagaagtaactacagtggtgggaaaaagatgcttagcact 2340
Db 2451 gaagaaattactcctctcttaagaagtaactacagtggtgggaaaaagatgcttagcact 2510
Qy 2341 ccaggagaatcaaaaattaaatttgacatggaaaaaatgcactcggctgttgggcaaggt 2400
Db 2511 ccaggagaatcaaaaattaaatttgacatggaaaaaatgcactcggctgttgggcaaggt 2570
Qy 2401 gtgcacgtctatccagagtgaaatctggaaattcttagctgagcaattccacactaaa 2460
Db 2571 gtgcacgtctatccagagtgaaatctggaaattcttagctgagcaattccacactaaa 2630
Qy 2461 caccagtttccagcaacagcagccaaagatgtgcatatacaagaactcttaagcag 2520
Db 2631 caccagtttccagcaacagcagccaaagatgtgcatatacaagaactcttaagcag 2690
Qy 2521 ctgacttcccgagcatgcatattcttattgaccttggcgaaaccttctcacacacca 2580
Db 2691 ctgacttcccgagcatgcatattcttattgaccttggcgaaaccttctcacacacca 2750
Qy 2581 tacttctctgccagcttggagcagcagcatctgcctttacaacattttgaaggcctac 2640
Db 2751 tacttctctgccagcttggagcagcagcatctgcctttacaacattttgaaggcctac 2810
Qy 2641 tcaacttctagaccaggaagtggatattgccaaggtctcagcttcttagcagggaatttg 2700
Db 2811 tcaacttctagaccaggaagtggatattgccaaggtctcagcttcttagcagggaatttg 2870

Qy 2701 cttcttcataatgagtgagggaagagcgctttaaaatgctcaagttcttgatatttgacatg 2760
Db 2871 cttcttcataatgagtgagggaagagcgctttaaaatgctcaagttcttgatatttgacatg 2930
Qy 2761 gggctgcggaaaaacagtagtcggccagacatgattattttacagatccagatgtaccagctc 2820
Db 2931 gggctgcggaaaaacagtagtcggccagacatgattattttacagatccagatgtaccagctc 2990
Qy 2821 tcgaggttgccttcatattaccacagagacctctacaatccctggaggagacagatc 2880
Db 2991 tcgaggttgccttcatattaccacagagacctctacaatccctggaggagacagatc 3050
Qy 2881 ggcceagcctcagctgccccctggttctccacctgtttgctccacagttcccgctg 2940
Db 3051 ggcceagcctcagctgccccctggttctccacctgtttgctccacagttcccgctg 3110
Qy 2941 ggaatcgtagcagagctctttgatatttttccagggaacagaggtcatattttaa 3000
Db 3111 ggaatcgtagcagagctctttgatatttttccagggaacagaggtcatattttaa 3170
Qy 3001 gtggtcttaagctctgttgggaagccataagcccttgattctgcagcatgaaaacctgaa 3060
Db 3171 gtggtcttaagctctgttgggaagccataagcccttgattctgcagcatgaaaacctgaa 3230
Qy 3061 acctagttgactttataaaaagcacgctacccaaaccttggcttgggtacagatgaaaag 3120
Db 3231 acctagttgactttataaaaagcacgctacccaaaccttggcttgggtacagatgaaaag 3290
Qy 3121 acctaatcaggtatttgaattggaatcgtaacagttcaaacagttacaagctttagaagttag 3180
Db 3291 acctaatcaggtatttgaattggaatcgtaacagttcaaacagttacaagctttagaagttag 3350
Qy 3181 taccacgtccttcaagaagaacttatcgattctctctctctcagtgacaacccaagaagt 3240
Db 3351 taccacgtccttcaagaagaacttatcgattctctctctctcagtgacaacccaagaagt 3410
Qy 3241 gataaattagagaacaaacacagcagcttacgcaaacagaaaccttgacctcttgaacag 3300
Db 3411 gataaattagagaacaaacacagcagcttacgcaaacagaaaccttgacctcttgaacag 3470
Qy 3301 ttgcaggtggcaaatggttagatccaaagccttgaggccaccattgagaagctcctgagc 3360
Db 3471 ttgcaggtggcaaatggttagatccaaagccttgaggccaccattgagaagctcctgagc 3530
Qy 3361 agtgagagcaagctgaagcagggccttaccttagaactggagcggtcgccctgctg 3420
Db 3531 agtgagagcaagctgaagcagggccttaccttagaactggagcggtcgccctgctg 3590
Qy 3421 cagacgttgagagagctcgggcgagcgagcgagccagccagccggagcctgagtg 3480
Db 3591 cagacgttgagagagctcgggcgagcgagcgagccagccagccggagcctgagtg 3650
Qy 3481 acgagcccgagcccgagcgagctgacagctctgcagggagagatttgcacacccatccca 3540
Db 3651 acgagcccgagcccgagcgagctgacagctctgcagggagagatttgcacacccatccca 3710
Qy 3541 cactgtccagggcctt 3555
Db 3711 cactgtccagggcctt 3725

RESULT 2

AAZ50907 standard; cDNA; 3988 BP.

XX AAZ50907;

XX AC (first entry)

XX DT 31-MAY-2000 (first entry)

XX DE Human TBC-1 cDNA from second transcript.

XX KW TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;

Db 1376 aggatagagggaatgaattcttccaaaacaaactagaaactgcaaaagccactgacagaca 1435
Qy 1261 ttaaccaatcaggagcaggaggaactatttttgaagagggttcagaatctgacacagagaat 1320
Db 1436 ttaaccaatcaggagcaggaggaactatttttgaagagggttcagaatctgacacagagaat 1495
Qy 1321 gagcagcagagagaatgaattgattattcttcttcttctgagatgtttatatgaagagaacag 1380
Db 1496 gagcagcagagagaatgaattgattattcttcttcttctgagatgtttatatgaagagaacag 1555
Qy 1381 aagaacacacatcccatattggggagatgaagcagacacacgagatggcagcagagaaattt 1440
Db 1556 aagaacacacatcccatattggggagatgaagcagacacacgagatggcagcagagaaattt 1615
Qy 1441 ggaagtgaattaccacccagctgcacactcgatttaggttagatatactgctgaaacacaaagca 1500
Db 1616 ggaagtgaattaccacccagctgcacactcgatttaggttagatatactgctgaaacacaaagca 1675
Qy 1501 aagagatctttaacagagcttttagaagatttttgcgggggttaataaagccagaggc 1560
Db 1676 aagagatctttaacagagcttttagaagatttttgcgggggttaataaagccagaggc 1735
Qy 1561 ctgcaggaaacactccatcagctggatctgtgatagcctccctgtctagtagtaactaaagtaac 1620
Db 1736 ctgcaggaaacactccatcagctggatctgtgatagcctccctgtctagtagtaactaaagtaac 1795
Qy 1621 accagcaaaagccatctgtgtgtgaaagaggcccttgccactctctgagagctcctt 1680
Db 1796 accagcaaaagccatctgtgtgtgaaagaggcccttgccactctctgagagctcctt 1855
Qy 1681 aagctctcggctcctcgaggaaactctccagtgactcggagagatctatcccaagaagag 1740
Db 1856 aagctctcggctcctcgaggaaactctccagtgactcggagagatctatcccaagaagag 1915
Qy 1741 ccagctccgctgtgcgcccagcagccttcaggaggcgagcaaacacctgagtcacttc 1800
Db 1916 ccagctccgctgtgcgcccagcagccttcaggaggcgagcaaacacctgagtcacttc 1975
Qy 1801 cccatgaatgccaggaaactccacaacctgcgcccgggttcccggggttttcgcaagg 1860
Db 1976 cccatgaatgccaggaaactccacaacctgcgcccgggttcccggggttttcgcaagg 2035
Qy 1861 aaacttatgaggtatcactcagtgagcacagagacgctcatgaacaaaggactttgaa 1920
Db 2036 aaacttatgaggtatcactcagtgagcacagagacgctcatgaacaaaggactttgaa 2095
Qy 1921 tccaaagcaaaacctcttgggtgattctgggtgggactcctgtgaagaccggaggcattcc 1980
Db 2096 tccaaagcaaaacctcttgggtgattctgggtgggactcctgtgaagaccggaggcattcc 2155
Qy 1981 tggaggcagcagatattcctccagtagtagcaccaccccgagagagggtcgatttctccagc 2040
Db 2156 tggaggcagcagatattcctccagtagtagcaccaccccgagagagggtcgatttctccagc 2215
Qy 2041 agatagaagattattcagagctgggagagcttccccacagatctctttagaacaggtt 2100
Db 2216 agatagaagattattcagagctgggagagcttccccacagatctctttagaacaggtt 2275
Qy 2101 tgtgaagatgggccccttggcccaccacagagagaaagaaagagacatctcgtgagctc 2160
Db 2276 tgtgaagatgggccccttggcccaccacagagagaaagaaagagacatctcgtgagctc 2335
Qy 2161 cgagagctgtggcaaaaggctattcttcaacagatactgctgttagaattggagaagaa 2220
Db 2336 cgagagctgtggcaaaaggctattcttcaacagatactgctgttagaattggagaagaa 2395
Qy 2221 aatcagagctccagcctctgaaatgatttctgaaacagccttgaagctcgattat 2280
Db 2396 aatcagagctccagcctctgaaatgatttctgaaacagccttgaagctcgattat 2455
Qy 2281 gaagaaataactcctgtctttaaagaagtaactacagtggtgggaaaaagatgcttagcact 2340
Db 2456 gaagaaataactcctgtctttaaagaagtaactacagtggtgggaaaaagatgcttagcact 2515

Qy 2341 ccaggaaagatacaaaaataagatttgacatggaaaaataatgcactcggtgttgggcaaggt 2400
Db 2516 ccaggaaagatacaaaaataagatttgacatggaaaaataatgcactcggtgttgggcaaggt 2575
Qy 2401 gtgcccagctcatcacccaggtgaaatctgaaattttctagctgagcaattccaccttaaa 2460
Db 2576 gtgcccagctcatcacccaggtgaaatctgaaattttctagctgagcaattccaccttaaa 2635
Qy 2461 caccagtttcccagcaaacacagcagccaaaggatgccaatacaagaactctttaaagcag 2520
Db 2636 caccagtttcccagcaaacacagcagccaaaggatgccaatacaagaactctttaaagcag 2695
Qy 2521 ctgacttcccagcagcagctgcgattcttattgacotttgggggaaaccttctctacacacca 2580
Db 2696 ctgacttcccagcagcagctgcgattcttattgacotttgggggaaaccttctctacacacca 2755
Qy 2581 tactctctcccagcagctgcgagcagcagctatcgctttacaacattttgaagccctac 2640
Db 2756 tactctctcccagcagctgcgagcagcagctatcgctttacaacattttgaagccctac 2815
Qy 2641 tcacttctagaccaggaagtgggatatgtccaaggtctcagcttttagcagggcattttg 2700
Db 2816 tcacttctagaccaggaagtgggatatgtccaaggtctcagcttttagcagggcattttg 2875
Qy 2701 cttcttcatatgagtgagggaaggcgctttaaaatgctcaagtttcttgatgtttgacatg 2760
Db 2876 cttcttcatatgagtgagggaaggcgctttaaaatgctcaagtttcttgatgtttgacatg 2935
Qy 2761 gggctgcgaaaaacagtagtcggccagacatgattattttacagatccagatgtaccagctc 2820
Db 2936 gggctgcgaaaaacagtagtcggccagacatgattattttacagatccagatgtaccagctc 2995
Qy 2821 tcgaggttgcttcatgattaccacagagacctctacaatcacctggggaggagcagagatc 2880
Db 2996 tcgaggttgcttcatgattaccacagagacctctacaatcacctggggaggagcagagatc 3055
Qy 2881 ggccccagcctctacgtgcctccctgggttctccacctgttggctcacagttcccgctg 2940
Db 3056 ggccccagcctctacgtgcctccctgggttctccacctgttggctcacagttcccgctg 3115
Qy 2941 ggattctagccagagctcttgatgatttttcttcagggaacagaggtctatatataa 3000
Db 3116 ggattctagccagagctcttgatgatttttcttcagggaacagaggtctatatataa 3175
Qy 3001 gtggctttaagctctgttgggaagccataagcccttgattctgcagcatgaaaaacctgaa 3060
Db 3176 gtggctttaagctctgttgggaagccataagcccttgattctgcagcatgaaaaacctgaa 3235
Qy 3061 accatagttgactttataaaaagcacgctaccacacctggcttgggtacagatggaaaag 3120
Db 3236 accatagttgactttataaaaagcacgctaccacacctggcttgggtacagatggaaaag 3295
Qy 3121 accatcaatcaggtattttgaaatggacatcgctaaacagttacaaagcttataagagttgag 3180
Db 3296 accatcaatcaggtattttgaaatggacatcgctaaacagttacaaagcttataagagttgag 3355
Qy 3181 taccacgtcttcaagaagaacttatcgattcctctctctcagtgacaaaccaagaagt 3240
Db 3356 taccacgtcttcaagaagaacttatcgattcctctctctcagtgacaaaccaagaagt 3415
Qy 3241 gataaaattagaaaaaaccaacagcagcttacgcaaacagacaccttgacctcttgaaacag 3300
Db 3416 gataaaattagaaaaaaccaacagcagcttacgcaaacagacaccttgacctcttgaaacag 3475
Qy 3301 ttgcaggtggcaaatggttaggattccaaagccttgaggccacctttagaagctcctgagc 3360
Db 3476 ttgcaggtggcaaatggttaggattccaaagccttgaggccacctttagaagctcctgagc 3535
Qy 3361 agtgagagcagctgaagcagggccatgcttaccttagaactggagcgggtcgccctgctg 3420
Db 3536 agtgagagcagctgaagcagggccatgcttaccttagaactggagcgggtcgccctgctg 3595

Db 943 cagatggcagcagagcattatggaaagtgaattaccaccacagtcgccactcgatttaggcta 1002
QY 1480 gatattgtgaaacaaagcaagagagatctttaacagagtccttttagaaagtattttgtcc 1539
Db 1003 gatattgtgaaacaaagcaagagagatctttaacagagtccttttagaaagtattttgtcc 1062
QY 1540 cgggttaataaagccagaggcctgcaggaacactccatcagtgtagctgtagatgctcc 1599
Db 1063 cgggttaataaagccagaggcctgcaggaacactccatcagtgtagctgtagatgctcc 1122
QY 1600 ctgctagtacattaaagcaacccagcaagagccatctgtgtgaaagaggcccttg 1659
Db 1123 ctgctagtacattaaagcaacccagcaagagccatctgtgtgaaagaggcccttg 1182
QY 1660 cccatctctgagagctctttaagctcctcgctcctcgaggacctgtccagtgactcg 1719
Db 1183 cccatctctgagagctctttaagctcctcgctcctcgaggacctgtccagtgactcg 1242
QY 1720 gagagtcattcccaagaagccagctccgctgtgcgcccagcagcccttcaggaggcga 1779
Db 1243 gagagtcattcccaagaagccagctccgctgtgcgcccagcagcccttcaggaggcga 1302
QY 1780 gcaaacacctgagtcacttcccacatcgaatgccaggaacctcccaacctgcccgggg 1839
Db 1303 gcaaacacctgagtcacttcccacatcgaatgccaggaacctcccaacctgcccgggg 1362
QY 1840 tccccgggggttcgaaaggaaacttatgaggtatcactcaagtgcagcacagagacgct 1899
Db 1363 tccccgggggttcgaaaggaaacttatgaggtatcactcaagtgcagcacagagacgct 1422
QY 1900 catgaacgaaggaacttgaatccaaagcaaacctatctggtgattctggtgggactcct 1959
Db 1423 catgaacgaaggaacttgaatccaaagcaaacctatctggtgattctggtgggactcct 1482
QY 1960 gtgaagccgggagcattctctggaggcagcagatattctcccgagtagccaccgccgag 2019
Db 1483 gtgaagccgggagcattctctggaggcagcagatattctcccgagtagccaccgccgag 1542
QY 2020 aaggcgtgcgattctccagcagatatgaagattattcagagctgggagagcttccccca 2079
Db 1543 aaggcgtgcgattctccagcagatatgaagattattcagagctgggagagcttccccca 1602
QY 2080 cgatctcttagaacagcttctgaagatggcccttggccccccaccagagaaag 2139
Db 1603 cgatctcttagaacagcttctgaagatggcccttggccccccaccagagaaag 1662
QY 2140 aaaaaggacatctcgtgagctcgagagctgtggcaaaaggctattcttcaacagatactg 2199
Db 1663 aaaaaggacatctcgtgagctcgagagctgtggcaaaaggctattcttcaacagatactg 1722
QY 2200 ctgcttagaattggaaggaadaatcagaagctccaaagcctctgaaaaatgatttgtgaac 2259
Db 1723 ctgcttagaattggaaggaadaatcagaagctccaaagcctctgaaaaatgatttgtgaac 1782
QY 2260 aagcgcctgaagctcgattatgaagaaatctactcctgtcttaaaagaagtaactacagtg 2319
Db 1783 aagcgcctgaagctcgattatgaagaaatctactcctgtcttaaaagaagtaactacagtg 1842
QY 2320 tgggaaaaagatcttagcactccaggaagatcaaaaaatgaatttgacatggaaaaaatg 2379
Db 1843 tgggaaaaagatcttagcactccaggaagatcaaaaaatgaatttgacatggaaaaaatg 1902
QY 2380 cactcgctgttggcaaggtgtccacgtcatcaccgaggtgaaatcttggaatttcta 2439
Db 1903 cactcgctgttggcaaggtgtgcaacgtcatcaccgaggtgaaatcttggaatttcta 1962
QY 2440 gctgagcaattccaccttaaacacagtttcccagcaaacagcagccaaagatgtgcca 2499
Db 1963 gctgagcaattccaccttaaacacagtttcccagcaaacagcagccaaagatgtgcca 2022
QY 2500 tacaagaactcttaagcagctgacttcccagcagcatgcgattcttattgaccttggg 2559
|||||

Db 2023 tacaagaactcttaagcagctgacttcccagcagcatgcgattcttattgaccttggg 2082
QY 2560 cgaaccttcttacacacaccatacttctctgcccagcttgagcaggagcagctacgctt 2619
Db 2083 cgaaccttcttacacacaccatacttctctgcccagcttgagcaggagcagctacgctt 2142
QY 2620 tacaacatttgaaggcctactcactcttagaccaggaagtgggatatatggcgaagttctc 2679
Db 2143 tacaacatttgaaggcctactcactcttagaccaggaagtgggatatatggcgaagttctc 2202
QY 2680 agctttgtagcaggcatcttcttctcatatgagtggaggaagcgctttaaaatgctc 2739
Db 2203 agctttgtagcaggcatcttcttctcatatgagtggaggaagcgctttaaaatgctc 2262
QY 2740 aagttcttgatgtttgacatgggctgcgaaaaacagtatcccgccagacagatatttita 2799
Db 2263 aagttcttgatgtttgacatgggctgcgaaaaacagtatcccgccagacagatatttita 2322
QY 2800 cagatccagatgtaccagctctcagaggtgtcttcagtgattaccacagagacctcacaa 2859
Db 2323 cagatccagatgtaccagctctcagaggtgtcttcagtgattaccacagagacctcacaa 2382
QY 2860 cacttgaggagcagcagatcgccccagcctctacgctgccccctggttctctcaccatg 2919
Db 2383 cacttgaggagcagcagatcgccccagcctctacgctgccccctggttctctcaccatg 2442
QY 2920 tttgctcacagttcccgctgggattctgtagccagagcttttgatatatttttctcag 2979
Db 2443 tttgctcacagttcccgctgggattctgtagccagagcttttgatatatttttctcag 2502
QY 2980 ggaacagaggtcattatttaaagtggtcttaagtcgttgggaagccataaagcccttgatt 3039
Db 2503 ggaacagaggtcattatttaaagtggtcttaagtcgttgggaagccataaagcccttgatt 2562
QY 3040 ctgcagcatgaaaacctagaaccatagttgactttataaaaaagcacgtaccaccaacct 3099
Db 2563 ctgcagcatgaaaacctagaaccatagttgactttataaaaaagcacgtaccaccaacct 2622
QY 3100 ggcctgtgacagatggaagaaccatcaatcaggtatttgaaatggacatcgctaaacag 3159
Db 2623 ggcctgtgacagatggaagaaccatcaatcaggtatttgaaatggacatcgctaaacag 2682
QY 3160 ttacaagcttatgaagttgagtcaccacgtctcttcaagaagaacttctgattctctctct 3219
Db 2683 ttacaagcttatgaagttgagtcaccacgtctcttcaagaagaacttctgattctctctct 2742
QY 3220 ctcaagtacaacccaagaatggataaaatagagaaaaacacacagcagcttaacgaaaaag 3279
Db 2743 ctcaagtacaacccaagaatggataaaatagagaaaaacacacagcagcttaacgaaaaag 2802
QY 3280 aaccttgacctcttgaacagttgcaggtggcaaatgtaggatccaaagccttgaggcc 3339
Db 2803 aaccttgacctcttgaacagttgcaggtggcaaatgtaggatccaaagccttgaggcc 2862
QY 3340 accattggaagctcctgagcagtgagagcaagctggaagcaggccatgcttaccctagaa 3399
Db 2863 accattggaagctcctgagcagtgagagcaagctggaagcaggccatgcttaccctagaa 2922
QY 3400 ctggagcaggtcggccctgctgcagacggtggagagctgcggcgcgagcagcagcc 3459
Db 2923 ctggagcaggtcggccctgctgcagacggtggagagctgcggcgcgagcagcagcc 2982
QY 3460 agcgaccgggagcctgagtgacgcagcccgagccacggcgactgcagcgtctgcaggg 3519
Db 2983 agcgaccgggagcctgagtgacgcagcccgagccacggcgactgcagcgtctgcaggg 3042
QY 3520 agagattgcaaacaccatcccacactgtccaggcctt 3555
|||||
Db 3043 agagattgcaaacaccatcccacactgtccaggcctt 3078
|||||

RESULT 4
AAC77404

Qy	1552	gcaagggccctgcaggaaacactccatcagtgctggatctcgtagtacctctgtctagtaca	1611
Db	1024	gccagaggccctgcaggaaacactccatcagtgctggatctcgtagtacctctgtagtaca	1083
Qy	1612	ttaagtaaaccagcaagaagccaactctgtgtgaaaaggagccctgcccatctctgag	1671
Db	1084	ttaagtaaaccagcaagaagccaactctgtgtgaaaaggagccctgcccatctctgag	1143
Qy	1672	agtcctcttaagctctctcgctcctcgcgaggacctgtcccagtgactcggagagtcattc	1731
Db	1144	agtcctcttaagctctctcgctcctcgcgaggacctgtcccagtgactcggagagtcattc	1203
Qy	1732	ccaggaagcccgagctccgctgtgcgcccagcaggccttcaggaaggcagcaaacacctg	1791
Db	1204	ccaggaagcccgagctccgctgtgcgcccagcaggccttcaggaaggcagcaaacacctg	1263
Qy	1792	agtcacttcccatacgaatgccagaaacctccaaacctgcccgggggtccccgggggtt	1851
Db	1264	agtcacttcccatacgaatgccagaaacctccaaacctgcccgggggtccccgggggtt	1323
Qy	1852	tcgcaaaggaaaccttatcaggttatcactcagtgagcacagagacgcctcatgaacgaag	1911
Db	1324	tcgcaaaggaaaccttatcaggttatcactcagtgagcacagagacgcctcatgaacgaag	1383
Qy	1912	gaotttgaatccaaaacaaacatatcttggtgatcttgtgtgggaactctgtgaagaccgg	1971
Db	1384	gaotttgaatccaaaacaaacatatcttggtgatcttgtgtgggaactctgtgaagaccgg	1443
Qy	1972	aggcattcctggaggcagcagatatctccagctagcaccccgaccccgacagaggcgtgcg	2031
Db	1444	aggcattcctggaggcagcagatatctcccgagtagccaccccgacagaggcgtgcg	1503
Qy	2032	tcttcagcagatatgaagattatcagagctgggagagcttccccacagatctccttta	2091
Db	1504	tcttcagcagatatgaagattatcagagctgggagagcttccccacagatctccttta	1563
Qy	2092	gaaccagtttgtgaagatgggccccttggccccccacagaggaaaaagaaagacatct	2151
Db	1564	gaaccagtttgtgaagatgggccccttgg - cccccacagaggaaaaagaaagacatct	1622
Qy	2152	cgtgagctccgagagctgtggcaaaaggctattctcaacagatctgc-tgctagaat	2210
Db	1623	cgtgagctccgagagctgtggcaaaaggcta tctcaacagatctgc-tgctagaat	1682
Qy	2211	ggagaaaggaaaa tcagaagctccaagcctctgaaatgatitgtgaacaaagccctgaa	2270
Db	1683	ggagaaaggaaaa tcagaagctccaagcctctgaaatgatitgtgaacaaagccctgaa	1742
Qy	2271	gctcgattatgaagaaattacccctctgttetaaagaagtaactacagtgggaaaaagat	2330
Db	1743	gctcgattatgaagaaattactccctgtcttaaagaagtaactacagtgggaaaaagat	1802
Qy	2331	gcttagcactccaggaagatcaaaaattaagtttgacatgaaaaaatgacactgcgt	2390
Db	1803	gcttagcactccaggaagatcaaaaattaagtttgacatgaaaaaatgacactgcgt	1862
Qy	2391	tgggcaagggtgtgccagctcatcccgaggtgaaatctggaaatttctagctgagcaatt	2450
Db	1863	tgggcaagggtgtgccagctcatcccgaggtgaaatctggaaatttctagctgagcaatt	1922
Qy	2451	ccaccttaaacaccagtttcccagcaaacagcagccaaggatgtgccatacaagaact	2510
Db	1923	ccaccttaaacaccagtttcccagcaaacagcagccaaggatgtgccatacaagaact	1982
Qy	2511	cttaagcagctgaacttcccagcagcatgcgatttcttattgaactttgggcgaaccttcc	2570
Db	1983	cttaagcagctgaacttcccagcagcatgcgatttcttattgaactttgggcgaaccttcc	2042
Qy	2571	tacacaccatactcttgcacagcttgagcaggacagctatcgcttttacacattt	2630
Db	2043	tacacaccatactcttgcacagcttgagcaggacagctatcgcttttacacattt	2102
Qy	2631	qaaggcctactcactctagacaggaaagtggatattggcaaggtctcagctttgtagc	2690

Db	2103	gaagcctactcaactctctagacaggaaagtggatattgccaaggctcagctttgtacg	2162
Qy	2691	aggcattttgctcttcatactagtgagaaagagcggtttaaaatgctcaagattctgat	2750
Db	2163	aggcattttgctcttcatactagtgagaaagagcggtttaaaatgctcaagattttctgat	2222
Qy	2751	gtttgacatggggctcggaacagtatcgccagacatgattattttacagatccagat	2810
Db	2223	gtttgacatggggctcggaacagtatcgccagacatgattattttacagatccagat	2282
Qy	2811	gtaccagctctgagggttgcttcagattaccacagagacccctcacatcactcgaggga	2870
Db	2283	gtaccagctctgagggttgcttcagattaccacagagacccctcacatcactcgaggga	2342
Qy	2871	gcaagatcgggcccgagcctctacgctgccctcggttctccaccaagtgttgcctcaca	2930
Db	2343	gcaagatcgggcccgagcctctacgctgccctcggttctccaccaagtgttgcctcaca	2402
Qy	2931	gtcccgctgggattcgtgacagagctcttgatatgatttttctcaagggaacagaggt	2990
Db	2403	gtcccgctgggattcgtgacagagctcttgatatgatttttctcaagggaacagaggt	2462
Qy	2991	cataattaaagtggcttgaagtctgttggaagacataaagcccttgattctgcagcatga	3050
Db	2463	cataattaaagtggcttgaagtctgttggaagacataaagcccttgattctgcagcatga	2522
Qy	3051	aaacctgaagaccatagttgactttataaaaagcagctaccacaccttggcttgggtaca	3110
Db	2523	aaacctgaagaccatagttgactttataaaaagcagctaccacaccttggcttgggtaca	2582
Qy	3111	gatgaaagaccatacaatcaggatttgaaatgacacatcgctaaacagttcacagctta	3170
Db	2583	gatgaaagaccatacaatcaggatttgaaatgacacatcgctaaacagttcacagctta	2642
Qy	3171	tgaagtgtgattacacgctcctcaagaagaaacttatcgattcctctcctcagtgacaa	3230
Db	2643	tgaagtgtgattacacgctcctcaagaagaaacttatcgattcctctcctcagtgacaa	2702
Qy	3231	ccaagaatggataaattgagagaaaaccaagcagcgcttacgcaaacagaaccttgacct	3290
Db	2703	ccaagaatggataaattgagagaaaaccaagcagcgcttacgcaaacagaaccttgacct	2762
Qy	3291	ccttgaacagtgtcaggtggaacttgtaggtaccacaaagccttgagggccaccattgagaa	3350
Db	2763	ccttgaacagtgtcaggtggaacttgtaggtaccacaaagccttgagggccaccattgagaa	2822
Qy	3351	gctcctgagcagtgagacgaactgaagcagggccatgcttaccttagaactggagcggtc	3410
Db	2823	gctcctgagcagtgagacgaactgaagcagggccatgcttaccttagaactggagcggtc	2882
Qy	3411	ggccctgtcacacaggttgaggagactcgcgcgagcagcagacccagcgaccggga	3470
Db	2883	g - -cctgctgcagcgggtggaggagctgcggcgcgagcagcagacccagcgaccggga	2940
Qy	3471	gcctgagtgcacgagcccgagcccacggcgactgcagctctgcaggagagattgcaa	3530
Db	2941	gcctgagtgcacgagcccgagcccacggcgactgcagct - tgcaggagagattgcaa	2999
Qy	3531	cacatccacacactgtccaggcctt	3555
Db	3000	caccat - -cacactgtccaggcctt	3022

RESULT	5
AAI59640	
ID	AAI59640 standard; cDNA; 3281 BP.
XX	
XX	
AC	AAI59640;
XX	
XX	
DT	22-OCT-2001 (first entry)
XX	
XX	
DE	Human polynucleotide SEQ ID NO 3629.

QY 1782 aaacacctgaagtcacttcccacatcgatccaggaacctccacaaacctgcccggggtc 1841
DB 1261 aaacacctgaagtcacttcccacatcgatccaggaacctccacaaacctgcccggggtc 1320
QY 1842 cccgggggttccgaaaggaaacttatgaggtatcactcagtgagcacagagacgctca 1901
DB 1321 cccgggggttccgaaaggaaacttatgaggtatcactcagtgagcacagagacgctca 1380
QY 1902 tgaacgaaggactttgaatccaaagcaaacatcttgggtattcttgggtgggactctgt 1961
DB 1381 tgaacgaaggactttgaatccaaagcaaacatcttgggtattcttgggtgggactctgt 1440
QY 1962 gaagaccggaggtattcctggagcgacagatattctcagatgacccaccgagaa 2021
DB 1441 gaagaccggaggtattcctggagcgacagatattctcagatgacccaccgagaa 1500
QY 2022 ggcgtcgattcttccagcagatatgaagattattcagagctggagagcttcccaccg 2081
DB 1501 ggcgtcgattcttccagcagatatgaagattattcagagctggagagcttcccaccg 1560
QY 2082 atctctttagaacacgtttgtgaagatgggccccttggccccccacacagaggaagaa 2141
DB 1561 atctctttagaacacgtttgtgaagatgggccccttggccccccacacagaggaagaa 1620
QY 2142 aagacatctcgtgagctccgagagctgtgcaaaagctattcttcaacagatactgct 2201
DB 1621 aagacatctcgtgagctccgagagctgtgcaaaagctattcttcaacagatactgct 1680
QY 2202 gcttagaattgagaggaataatcagaagctccagagctctgaaatgatttgcgaaaca 2261
DB 1681 gcttagaattgagaggaataatcagaagctccagagctctgaaatgatttgcgaaaca 1740
QY 2262 ggcctgaagctcgtattatgaagaattactcctcgtcttaagaagataactacagtgct 2321
DB 1741 ggcctgaagctcgtattatgaagaattactcctcgtcttaagaagataactacagtgct 1800
QY 2322 ggaagaatgcttagcactcaggaagatcaaaaattgaatttgacatggaataatgca 2381
DB 1801 ggaagaatgcttagcactcaggaagatcaaaaattgaatttgacatggaataatgca 1860
QY 2382 ctgcgctgttgggaaggtgtg-ccagctcatccagaggtgaaatttggaattttctag 2440
DB 1861 ctgcgctgttgggaaggtgtgcccagctcatccagaggtgaaatttggaattttctag 1920
QY 2441 ctgagcaattccacattcaaacacagtttcccagcaaacagcagcgaagatgtgccat 2500
DB 1921 ctgagcaattccacattcaaacacagtttcccagcaaacagcagcgaagatgtgccat 1980
QY 2501 acaagaaactcttaagcagctgacttcccagcagcagcagcagcagcagcagcagcagc 2560
DB 1981 acaagaaactcttaagcagctgacttcccagcagcagcagcagcagcagcagcagcagc 2040
QY 2561 gaaccttctctac 2620
DB 2041 gaaccttctctac 2100
QY 2621 acaaatattgaaagcctactcacttcttagaccaggaagtggaattgccaaggtctca 2680
DB 2101 acaaatattgaaagcctactcacttcttagaccaggaagtggaattgccaaggtctca 2160
QY 2681 gctttgtagcaggaatttgccttctcatatgagtggaagagcgcttttaaaatgctca 2740
DB 2161 gctttgtagcaggaatttgccttctcatatgagtggaagagcgcttttaaaatgctca 2220
QY 2741 agttcttgatgtttgacatggggtcggaacacagatcgccagacatgattattttac 2800
DB 2221 agttcttgatgtttgacatggggtcggaacacagatcgccagacatgattattttac 2280
QY 2801 agatccagatgtaccagctctcaggttgccttcatgattaccacagagacctctacaatc 2860
DB 2281 agatccagatgtaccagctctcaggttgccttcatgattaccacagagacctctacaatc 2340

QY 2861 acctggaggagcagcagatcgccccagcctctacgctgccccctgggttctcaccatgt 2920
DB 2341 acctggaggagcagcagatcgccccagcctctacgctgccccctgggttctcaccatgt 2400
QY 2921 ttgctcacagttcccgcgtgggattcgtagccagagtcctttgatgatattttttcagg 2980
DB 2401 ttgctcacagttcccgcgtgggattcgtagccagagtcctttgatgatattttttcagg 2460
QY 2981 gaacagaggtcatttttaaaagtggctttaagtctgttgggaagcacaataagccctgattc 3040
DB 2461 gaacagaggtcatttttaaaagtggctttaagtctgttgggaagcacaataagccctgattc 2520
QY 3041 tgacagatgaaacctgaaagaccatagttagctttataaaagcagctacccaaaccttg 3100
DB 2521 tgacagatgaaacctgaaagaccatagttagctttataaaagcagctacccaaaccttg 2580
QY 3101 gcttggtacagatgaaagaccatcaatcaggtatttgaaatggacatcgctaacaagt 3160
DB 2581 gcttggtacagatgaaagaccatcaatcaggtatttgaaatggacatcgctaacaagt 2640
QY 3161 tacaagtttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctctc 3220
DB 2641 tacaagtttatgaagttgagtaccacgtccttcaagaagaacttatcgattcctctctc 2700
QY 3221 tcagtgacaaccaaagaatggataaaattagagaaaccaaagcagcttacgcaaacaga 3280
DB 2701 tcagtgacaaccaaagaatggataaaattagagaaaccaaagcagcttacgcaaacaga 2760
QY 3281 acctgacctctcttgaacagttgcaggtggcaaatggttaggatccaaagccttgaggcca 3340
DB 2761 acctgacctcttgaacagttgcaggtggcaaatggttaggatccaaagccttgaggcca 2820
QY 3341 ccattgagaagcctcctgagcagtgagagcaagctggaagcagcagcctgttacccttagaac 3400
DB 2821 ccattgagaagcctcctgagcagtgagagcaagctggaagcagcagcctgttacccttagaac 2880
QY 3401 tgagcggctcgccctgctgcagacggtggagggagctgcgcgagcgagcgagcgcca 3460
DB 2881 tgagcggctcgccctgctgcagacggtggagggagctgcgcgagcgagcgagcgcca 2940
QY 3461 gcgaccggagcctgagtgacgcagcccgcagcccagcggcgagcctgtgcagga 3520
DB 2941 gcgaccggagcctgagtgacgcagcccgcagcccagcggcgagcctgtgcagga 3000
QY 3521 gagattgcaacacacatcccacactgtccaggcctt 3555
DB 3001 gagattgcaacacacatcccacactgtccaggcctt 3035

RESULT 6

AAV05886

ID AAV05886 standard; DNA; 4039 BP.

XX

AC AAV05886;

XX

DT 01-JUN-1998 (first entry)

XX

DE Human Tbc-1 gene.

XX

KW tre-2; BUB2; cdc16; mast cell; probe; hybridisation; subtraction method;
mouse; transcription factor; differentiation; proliferation; human; ds;

KW acute myelogenous leukaemia.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

CDS 403..3828

FT /*tag= a

FT /product= Tbc1_protein

XX

PN US5700927-A.

XX

PD 23-DEC-1997.

XX 23-DEC-1994; 94US-0363300.
XX 23-DEC-1994; 94US-0363300.
XX (CHIL-) CHILDRENS MEDICAL CENT.
XX Richardson P, Zon L;
XX WPI; 1998-062437/06.
XX P-PSDB; AAW44777.
XX DNA encoding Tbc1 polypeptide - useful for treating leukaemia
XX Claim 3; Fig 1A-B; 2pp; English.
XX This is the nucleotide sequence which encodes a Tbc1 (tre-2, BUB2 and
XX cdc16) polypeptide, isolated from a bone marrow-derived mast cell DNA
XX library. The screen was carried out using a probe generated by a
XX subtraction method which compared mRNA expression in an undifferentiated
XX mast cell line P815 and cell line PC76 (P815 cells transformed to express
XX the murine GATA-1 transcription factor - a factor which controls the
XX expression of genes involved in mast cell differentiation). Tbc1 encodes
XX a protein involved in the coupling of cell proliferation to cell
XX differentiation, which can be used to treat leukaemia (especially acute
XX myelogenous leukaemia) by causing leukaemic cells to differentiate.
XX Sequence 4039 BP; 1042 A; 1096 C; 1060 G; 841 T; 0 other;
SQ

Query Match 70.6%; Score 2510.4; DB 19; Length 4039;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 2949; Conservative 0; Mismatches 536; Indels 37; Gaps 8;

QY 1 atggagcaataacattcacaggaagaaacatctgctt-cttaacaggggtctcggtga 59
DB 292 atggaggcaatcacattcacaggaagaaacatctgcttctccataacgaagtcctcgga 351
QY 60 ttttgctgcagctggtt-ggctccctgctgctgcttccctgacacacacgctccatg 117
DB 352 ctctggctgcagctggttgggtctctcttaccctgctgcttctctcaccactgctccatg 411
QY 118 ctgctctgggttgggtgagtgaggaagactcagcagcagcagcagcagcagcagcagcagc 177
DB 412 ctgctgggttgggtgagtgaggaagactcagcagcagcagcagcagcagcagcagcagc 471
QY 178 gtaaccaagcaagtcggttgggttccacccctgctgactgagatgtaacctgagcca 237
DB 472 aggaacaagcaagtcggttgggttccacccctgctgactgagatgtaacctgagcca 531
QY 238 ggaagaagcaagtcggttgggttccacccctgctgactgagatgtaacctgagcca 297
DB 532 ggaagaagcaagtcggttgggttccacccctgctgactgagatgtaacctgagcca 591
QY 298 cgtgttcacaaactgattcacacagtcagcagcagcagcagcagcagcagcagcagcagc 357
DB 592 cgtgtgcacaaactgattcacacagtcagcagcagcagcagcagcagcagcagcagcagc 651
QY 358 gaagacgctgtccacgagcagagtcctgctatgttttcaagcgcagatgatacaacaaaa 417
DB 652 gaggatgctgcacagcagagcctctgctatgttttcaagcgcagatgatacaacaaaa 711
QY 418 gtgcctgagatcatcagctccatcctgagcggggaagatcgcccgagcagagagctg 477
DB 712 gtgcctgagatcatcagctccatcctgagcggggaagatcgcccgagcagagagctg 771
QY 478 cactgcccctcgagttcagcagcagcttttccaaagtttcgaggtgctctctcgccg 537
DB 772 cgttgcctcctcgagttcagcagcagcttttccaaagtttcgaggtgctctctcgccg 831
QY 538 cgcgtgacgttggtgcacaaaggtctccgcccgcctgactgacgagtgcatcgagaag 597
DB 832 cgggtgactgtggtgcacaaaggtctccgcccgcctgactgacgagtgcatcgagaag 891

QY 598 ttcaatcagtcagcgagcgagcggttccgagagccccccgcccccaacccgccccatgcc 657
DB 892 ttcaaccatgtgagctgtggtcgcgagacggtggaagcgcccccagcggtgagcagcatca 951
QY 658 gcgcccacagggagcagagagcctgtgcgagggcccatgcgaagtctcttcccagccc 717
DB 952 gc-----gctggccccagggcccatgcgcaaatctcttcacagcct 993
QY 718 ggcctgcgctgctgctgctttaggaagagctgcagaggtggtggggtcccgagcagcggc 777
DB 994 ggaactgcgctgctgctgcttccaggaaggttccaggaagcgtagctcgcagtagtacc 1053
QY 778 ttctcagctccttcgagagagcgacattgagacacacctcattagcgagacacaaatt 837
DB 1054 tt-----tagctcttggacaatgacatagagaacacctcattcgtgtgggcacattgtg 1107
QY 838 gtgcagccacagatctcaggaagaaatcgaaactatgctcttcacagattggccagtgtaa 897
DB 1108 gtgcagccacagatctcaggaagaaatcgaaactatgctcttcacagattggccagtgtaa 1167
QY 898 gttacctcactcagtcctgcagacacacacacacacacacacacacacacacacacacac 957
DB 1168 gttacctcactcagtcctgcagacacacacacacacacacacacacacacacacacacac 1227
QY 958 tcttttctcctcagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1017
DB 1228 tcttttctcctcagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1287
QY 1018 tc-----cggaggtggcggttctcattttgtctgttactgttcttactgttactgttactgt 1071
DB 1288 tcgggtggcggtgagtggttctcattttgtctgttactgttcttactgttactgttactgtt 1347
QY 1072 gctcgtgttgatgaataatgatgacctgaaacagccttcacagctgcagctgcagctgcag 1131
DB 1348 gctcgtgttgatgaataatgatgacctgaaacagccttcacagctgcagctgcagctgcag 1407
QY 1132 cagacagctaaagcgccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1191
DB 1408 cagacagctaaagcgccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1467
QY 1192 ctctgtgagagatgagggagaaatcttccaaacacacacacacacacacacacacacacac 1251
DB 1468 ctctgtgagagatgagggagaaatcttccaaacacacacacacacacacacacacacacac 1527
QY 1252 ctgcagcaatcaacaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1311
DB 1528 ttgcacacactgacacaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1587
QY 1312 ccgaaatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1371
DB 1588 ccaagaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1647
QY 1372 ggaacacagaaagac 1431
DB 1648 gagaagcaaaagac 1707
QY 1432 gagaatattggaaggaattaccacccagtcgcaactcgtatttaggctagatatgctgaaa 1491
DB 1708 gagaatattggaaggaattaccacccagtcgcaactcgtatttaggctagatatgctgaaa 1767
QY 1492 aacaagcaaaagagatctttaaagagctctttaaagatatttttcccggtggtataaaa 1551
DB 1768 aacagcaaaagagatctttaaagagctctttaaagatatttttcccggtggtataaaa 1827
QY 1552 gccagagcctgcagaaac 1611
DB 1828 gccagagcctgcagaaac 1887
QY 1612 ttaagtaac 1671
DB 1888 ctaagtaac 1947

Qy	1672	agctctttaagctcctctcgccctctctggaggagacctgttccagtgtactgcggagatcatctc	1731
Db	1948	acctctccaagctctcttgctctctcagatgacctgtccagtgtcagagggccacatt	2007
Qy	1732	ccagaagagccagctccgctgtctgcccagcagccttcaggagcagcaaacaccctg	1791
Db	2008	gcagaagagctctgccctgtgtcacccagcaggcgttcagaagagagccaacaccctg	2067
Qy	1792	agtcacttcccactcgaaatgccaggaacctccacaacctgcccgccgggggtccccgggggtt	1851
Db	2068	agtcatttcccagtagagtgccctgcgctccagaacctgccagagctctccaggggtc	2127
Qy	1852	tcgcaaggaaacctatgaggctatcactcaatgtagcacagacagcgtctatgaacgaag	1911
Db	2128	tctcaaggaaacctatgagttaccactccgtgagcacagagcgtctatgaacgcaag	2187
Qy	1912	gactttgaatccaaagcaaaccaattcttgatctctggtggacctctgtgaagaccgg	1971
Db	2188	gactttgaa tccaaagcaaaccaacctggtgacacagatgggaccccgctgaagaccgg	2247
Qy	1972	aggcattctctggaggcagcagatattcttcgagtagccaccccgcaagagggtgcgat	2031
Db	2248	cggcactctggagacagcagatattctctgagtgccactccacagaaggcttgtgac	2307
Qy	2032	tcttcagcagatatgaagattattcagagctgggagagcttccccacagatctcctta	2091
Db	2308	tcgccagcagatatgaagattattccgagctgggagagctcctccacgctccccctta	2367
Qy	2092	gaaccagttgtgaagatggggccctttggccccccacacagaggagaaaaagagacatct	2151
Db	2368	gaaccgggtgtgaggcggccactttgg - ccagtacaggagaaaaagagagaagctca	2426
Qy	2152	cgtgagctccgagagctgtggcaaaagcctattctcaacagatactg - ctgcttagaat	2210
Db	2427	cggagcttcgagagctgtggaaaaaggccactctgcagcagatccttgctctcgaggat	2486
Qy	2211	ggagaggaataatcagaagctccaagctctgaataatgattgtctgaacagcgctgaa	2270
Db	2487	ggagagaggaaatcagaagctcaagcctctgaatacgtattgtctgaacaaagcctcaa	2546
Qy	2271	gctcgattatgaagaataattactccctgtcttaagaagaactaactacagtgaggaaaaagt	2330
Db	2547	gcttgacta tgaagaataacacccgtgtcttaagaagctcactacagtgaggaaaaagt	2606
Qy	2331	gcttagcactccaggaagatcaaaataaagtttgacatggaaaaaatgcactcggctgt	2390
Db	2607	gcttagcactccaggaagatccaaaataaagtttgacatggaaaaagtgcactcagctgt	2666
Qy	2391	tgggcaagg tgtgccagctca tcccgagg tgaatacttggaaattcttagctgagcaatt	2450
Db	2667	tgggcaagg tgtgccagctca tcccgagg tgaatctggaatttctagctgagcagtt	2726
Qy	2451	ccacctaaacacagatttcccagcaaacacagcagccaagagatgtgccataaagaact	2510
Db	2727	ccaccttaaacacccatttccctagttaaacagcagccaagagcgtgccctacaaagact	2786
Qy	2511	cttaaaagcagctgacttcccagcagcatgcgattcttattgaccttggggcgaaactttcc	2570
Db	2787	ccgtgaagaagctgaactctcgacgcagcaagccattctcatgacctcggggcgaaactttcc	2846
Qy	2571	tacacacccatactctctgccagcttggagcagcagagctatcgctttacaacatttt	2630
Db	2847	aacacatccatacttctctgccagcttggagcaggtcagctgtcactttacaacattct	2906
Qy	2631	gaaggcctactcactcttagacacagaagtggaattatggcaaggctctcagctttgtagc	2690
Db	2907	gaaggcctactcgtcttgagccagaggttgga tacttgccaaggtctcagctttgtggc	2966
Qy	2691	agccatttgcctctctcatatgagtgaggaagagcgctttaaaatgctcgaattctcgat	2750
Db	2967	agggatttgcctctctcaatagtgaggaagggcgctcgaagtgtccaagttccctgat	3026
Qy	2751	qtttgacatagggctgcgaataacagtatctcgccacacagatgattatttacaagatccagat	2810

[illegible]

RESULT 7

AAH18401
ID AAH18401 standard; cDNA; 2362 BP.

AAH18401:

DT 26-JUN-2001 (first entry)

XX
DE Human cDNA sequence SEQ ID NO:18465

XX Human: primer: detection: diagnosis: antisense therapy; gene therapy; ss.

XX
Homo sapiensXX
PN
FD1074617-A2XX
07 FEB - 2001

XX
XX
XX

00000-0116136

XX

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 18465; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-qt primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 2362 BP; 657 A; 581 C; 578 G; 546 T; 0 other;
SQ

Query Match 59.4%; Score 2112; DB 22; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1444 agtgaattaccacccagtcgactcgtattagctagatgctgaaacaaagcaag 1503
Db 1 agtgaattaccacccagtcgactcgtattagctagatgctgaaacaaagcaag 60

Qy 1504 agatctttaacagagcttttagaagattttgtcccggttaataaagcagggcgt 1563
Db 61 agatctttaacagagcttttagaagattttgtcccggttaataaagcagggcgt 120

Qy 1564 caggaaacacccatcagtggtgactcgtagctccctgtctagtaataagtaacacc 1623
Db 121 caggaaacacccatcagtggtgactcgtagctccctgtctagtaataagtaacacc 180

Qy 1624 agcaaaagccatctgtgtgaaagaggccttgcccatctctgagactcctttaag 1683
Db 181 agcaaaagccatctgtgtgaaagaggccttgcccatctctgagactcctttaag 240

Qy 1684 ctctcgtcctcctcggaggacctgtccagtgactcgtgagagtaactcaccagaagacca 1743
Db 241 ctctcgtcctcctcggaggacctgtccagtgactcgtgagagtaactcaccagaagacca 300

Qy 1744 gctccgtgtgcgccagagccttcagaggcagcacaacacctgaactctccc 1803
Db 301 gctccgtgtgcgccagagccttcagaggcagcacaacacctgaactctccc 360

Qy 1804 atcgaatgcaggaacacctccacaaacctgcccgggggtccccgggggttttcgcaaaagaaa 1863
Db 361 atcgaatgcaggaacacctccacaaacctgcccgggggtccccgggggttttcgcaaaagaaa 420

Qy 1864 cttatgaggtatcactcagtgagcagagacgcctcctaatgaacgaagagactttgaaatcc 1923
Db 421 cttatgaggtatcactcagtgagcagagacgcctcctaatgaacgaagagactttgaaatcc 480

Qy 1924 aaagcaaaacatcttggtgattctggtgagctccctctgtaagaccgagagcattctctg 1983
Db 481 aaagcaaaacatcttggtgattctggtgagctccctctgtaagaccgagagcattctctg 540

Qy 1984 aggcagcagatattctccgagtagccaccccgacagaagcggtgcgattcttccagcaga 2043
Db 541 aggcagcagatattctccgagtagccaccccgacagaagcggtgcgattcttccagcaga 600

Qy 2044 tatgaagattattcagagctggagagcttcccccagcatctcttttagaacagatttgt 2103
Db 601 tatgaagattattcagagctggagagcttcccccagcatctcttttagaacagatttgt 660

Qy 2104 gaagatgggcccctttggccccccacccagagagaaaagaaagagacatctcgtgagctccga 2163
Db 661 gaagatgggcccctttggccccccacccagagagaaaagaaagagacatctcgtgagctccga 720

Qy 2164 gagctgtggcaaaagcctattcttcaacagatgactgctgttagaatggagaagaaaaat 2223
Db 721 gagctgtggcaaaagcctattcttcaacagatgactgctgttagaatggagaagaaaaat 780

Qy 2224 cagaagctcaaacctctgaaatgatttgcgaacaaagcgcctgaagctcgattatgaa 2283
Db 781 cagaagctcaaacctctgaaatgatttgcgaacaaagcgcctgaagctcgattatgaa 840

Qy 2284 gaattactccctgtctttaaagaagtaactacagtggtgggaaaagatgcttagcactcca 2343
Db 841 gaattactccctgtctttaaagaagtaactacagtggtgggaaaagatgcttagcactcca 900

Qy 2344 ggaagatcaaaaattaaagtttgacatggaaaaaatgcactcggctgttggcgaagtggtg 2403
Db 901 ggaagatcaaaaattaaagtttgacatggaaaaaatgcactcggctgttggcgaagtggtg 960

Qy 2404 ccagctcatccacgaggtgaaatctgaaatttctagctgagcaattcccacaccttaaacac 2463
Db 961 ccagctcatccacgaggtgaaatctgaaatttctagctgagcaattcccacaccttaaacac 1020

Qy 2464 cagtttcccagcaaacagcagcgaagagtggtccatacaaaaactctttaaagcagctg 2523
Db 1021 cagtttcccagcaaacagcagcgaagagtggtccatacaaaaactctttaaagcagctg 1080

Qy 2524 acttcccagcagcatgcttatttattgaccttggcgaaaccttctcacaacccatcac 2583
Db 1081 acttcccagcagcatgcttatttattgaccttggcgaaaccttctcacaacccatcac 1140

Qy 2584 ttctctgccagcttgagcagcagcagctatcgtttacacatttttgaggcctactca 2643
Db 1141 ttctctgccagcttgagcagcagcagctatcgtttacacatttttgaggcctactca 1200

Qy 2644 ctctagacaggaagtgggatatattgccaaggtctcagctttgttagcagcattttgctt 2703
Db 1201 ctctagacaggaagtgggatatattgccaaggtctcagctttgttagcagcattttgctt 1260

Qy 2704 ctctcatagtgtaggaagaggcgttttaaaatgctcaagtttctgagtgatttgacatgggg 2763
Db 1261 ctctcatagtgtaggaagaggcgttttaaaatgctcaagtttctgagtgatttgacatgggg 1320

Qy 2764 ctgcgaaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 2823
Db 1321 ctgcgaaaacagtatcggccagacatgattattttacagatccagatgtaccagctctcg 1380

Qy 2824 aggtgtcttcatgatattaccacagagacctctacaatccctggaggagcagcagatcggc 2883
Db 1381 aggtgtcttcatgatattaccacagagacctctacaatccctggaggagcagcagatcggc 1440

Qy 2884 cccagcctctacgctgccccctgttctcctcaccatgttttgcctcacagttcccgcgtggga 2943


```
Db 1441 cccgctctacgtgccccctggttccctcaccatggttgcctcacagttcccgctggga 1500
Qy 2944 ttcgtaccagagctcttgcattatattttcttcagggaacagaggtcattattaaagt 3003
Db 1501 ttcgtaccagagctcttgcattatattttcttcagggaacagaggtcattattaaagt 1560
Qy 3004 gctttaagctctgttggaagccataaagcccttgattctgcagcatgaaacccctagaacc 3063
Db 1561 gctttaagctctgttggaagccataaagcccttgattctgcagcatgaaacccctagaacc 1620
Qy 3064 atagttgacttataaaagcagcgtaccaccccttgcttggttacagatggaagacc 3123
Db 1621 atagttgacttataaaagcagcgtaccaccccttgcttggttacagatggaagacc 1680
Qy 3124 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttgaagttgagtac 3183
Db 1681 atcaatcaggtatttgaaatggacatcgctaaacagttacaagcttgaagttgagtac 1740
Qy 3184 cagctctcaagaagaacttatcgattctctctctctcgtgacaaacccaaagaatggat 3243
Db 1741 cagctctcaagaagaacttatcgattctctctctcgtgacaaacccaaagaatggat 1800
Qy 3244 aaattagaaaaaacacagcagcttacgcaaacagaaacctgaacctcttgaaacagttg 3303
Db 1801 aaattagaaaaaacacagcagcttacgcaaacagaaacctgaacctcttgaaacagttg 1860
Qy 3304 caggtgcaaatgtagtattcctaaagccttgaggccaccattgagaagctcctgagcagt 3363
Db 1861 caggtgcaaatgtagtattcctaaagccttgaggccaccattgagaagctcctgagcagt 1920
Qy 3364 gagagcaagctgaagcaggcctgcttaactctgaaactgagcggctgctgctgcag 3423
Db 1921 gagagcaagctgaagcaggcctgcttaactctgaaactgagcggctgctgctgcag 1980
Qy 3424 acggtgagagctgcgcggcgagcgcagagcccgagcggcgagcctgagctgagtcagc 3483
Db 1981 acggtgagagctgcgcggcgagcgcagagcccgagcggcgagcctgagctgagtcagc 2040
Qy 3484 cagcccgagcccgagcggcgagcctgagcagctgagcagagattgcaacacccatccacac 3543
Db 2041 cagcccgagcccgagcggcgagcctgagcagctgagcagagattgcaacacccatccacac 2100
Qy 3544 tgtccaggcctt 3555
Db 2101 tgtccaggcctt 2112

RESULT 8
AAS34787
ID AAS34787 standard; cDNA; 1422 BP.
XX AC AAS34787;
XX XX
DT 04-DEC-2001 (first entry)
XX XX
DE cDNA encoding novel human neoplastic disease associated polypeptide #21.
XX XX
KW Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX OS Homo sapiens.
XX PN W0200155163-A1.
XX XX
PD 02-AUG-2001.
XX XX
PF 17-JAN-2001; 2001WO-US01358.
XX XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
```


PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI: 2001-465558/50.
 DR P-FSDB; AAU21586.
 XX
 XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
 PT diagnose diseases or disorders associated with aberrant expression or
 PT activity of polypeptides, and for treating cancers, rheumatoid
 PT arthritis
 XX
 PS Claim 4; SEQ ID No 29; 687pp; English.
 XX
 CC The present invention relates to the isolation of novel human neoplastic
 CC disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
 CC sequences encoding for these polypeptides. The sequences of the
 CC invention are useful in the diagnosis, treatment, prevention and/or
 CC prognosis of disorders involving neoplastic disease such as
 CC hyperproliferative disorders (e.g. leukemia, bone cancer, bladder
 CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
 CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
 CC also be useful for treating other disorders such as neural disorders,
 CC immune system disorders, muscular disorders, reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
 CC and renal disorders. The polynucleotide sequences of the invention are
 CC also useful in gene therapy. AAS34767-AAS35050 represent cDNA sequences
 CC encoding for the novel human neoplastic disease associated polypeptides
 CC of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1269 BP; 362 A; 313 C; 277 G; 312 T; 5 other;

 Query Match 26.5%; Score 943.2; DB 22; Length 1269;
 Best Local Similarity 98.7%; Pred. No. 7.7e-247;
 Matches 967; Conservative 3; Mismatches 8; Indels 2; Gaps 2;

 QY 1908 aaaggacttgaatccaagaacacacattcttggtgattcttggtgggactctgtgaaagac 1967
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 292 aagggaacttgaatccaagaacacacattcttggtgattcttggtgggactctgtgaaagac 351

 QY 1968 ccgagggaattcttgaggcagcagatattcttcgagtagcaccgccgcgagaagcgctg 2027
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 352 ccgagggaattcttgaggcagcagatattcttcgagtagcaccgccgcgagaagcgctg 411

 QY 2028 cgattctccagcagatgaagattattcagagctggagagcttccccacgactccc 2087
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 412 cgattctccagcagatgaagattattcagagctggagagcttccccacgactccc 471

RESULT 11

AAH04450

ID AAH04450 standard; cDNA; 883 BP.

XX AC AAH04450;

XX AC AAH04450;

DT 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:1285.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX

PD 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

PF 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

PI WPI; 2001-318749/34.

XX

XX Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

XX

XX Claim 1; SEQ ID 1285; 2537pp + CD ROM; English.

PS

XX The present invention describes primer sets for synthesising 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification

CC of the present invention.

XX

XX Sequence 883 BP; 245 A; 221 C; 219 G; 193 T; 5 other;

SQ

Query Match 22.1%; Score 785; DB 22; Length 883;

Best Local Similarity 96.7%; Pred. No. 9.9e-204;

Matches 843; Conservative 0; Mismatches 23; Indels 6; Gaps 4;

Qy 1444 agtgaattaccacccagtcgactcgtatttagctgagatgctgaaacaaagcaag 1503

Db 1 agtgaattaccacccagtcgactcgtatttagctgagatgctgaaacaaagcaag 60

Qy 1504 agatctttaacagagcttttagaagattttgtcccggtgtaataaagccagagcctg 1563

Db 61 agatctttaacagagcttttagaagattttgtcccggtgtaataaagccagagcctg 120

Qy 1564 caggaacacccatcagctgagatgctgagatgctcctgctcctagcattaaataacacc 1623

Db 121 caggaacacccatcagctgagatgctgagatgctcctgctcctagcattaaataacacc 180

Qy 1624 agcaaaagcgcattctgtgtaaaagagccttgccctctctcctagagctcctttaag 1683

Db 181 agcaaaagcgcattctgtgtaaaagagccttgccctctctcctagagctcctttaag 240

Qy 1684 ctctcggctcctcggaggaccctgtccagtgactcggagagtgactctccagagagccca 1743

Db 241 ctctcggctcctcggaggaccctgtccagtgactcggagagtgactctccagagagccca 300

Qy 1744 gctccgctgtcgcacccagcagcgttcagaggcgagcaaaacacccctgagtcactcccc 1803

Db 301 gctccgctgtcgcacccagcagcgttcagaggcgagcaaaacacccctgagtcactcccc 360

Qy 1804 atcgaatgcaggaacacctccacaacctgcccgggggtcccgggggttttcgcaaaagaaa 1863

Db 361 atcgaatgcaggaacacctccacaacctgcccgggggtcccgggggttttcgcaaaagaaa 420

Qy 1864 cttatgaggtatcactcagtgagcagacagacccctcatgaacaaagactttgaaatcc 1923

Db 421 cttatgaggtatcactcagtgagcagacagacccctcatgaacaaagactttgaaatcc 480

Qy 1924 aaagcaaacacctgtgtgactcctggtgagactcctgttaaaagcccgagagcattctcg 1983

Db 481 aaagcaaacacctgtgtgactcctggtgagactcctgttaaaagcccgagagcattctcg 540

Qy 1984 aggcagcagatatctcccgagtagccaccccgcaagaaggcgtgcgattctctccagcaga 2043

Db 541 aggcagcagatatctcccgagtagccaccccgcaagaaggcgtgcgattctctccagcaga 600

Qy 2044 tatgaagattattcagagctgggagagcttccccacgactctcctttagaacagtttgt 2103

Db 601 tatgaagattattcagagctgggagagcttccccacgactctcctttagaacagtttgt 659

Qy 2104 gaagatggccctttggccccccacccagagagaaagaaagacatctcgtgagctccga 2163

Db 660 gaagatggccctttggccccccacccagagagaaagaaagacatctcgtgagctccga 719

Qy 2164 gagc-tgtggcaaaagctattcttcaacagatactgctgctttagaataggagagaa 2222

Db 720 agcttgtggcaaaagctattcttcaacagatactgctgctttagaataggagagaa 779

Qy 2223 taagaagctccaagc---tctgaaaatgatttgcgtgaacaaagcgc-tgaagctcgatt 2278

Db 780 taagaagctccaagc---tctgaaaatgatttgcgtgaacaaagcgc-tgaagctcgatt 839

Qy 2279 atgaagaataactcctcctcctccttaagaagagta 2310

Db 840 ntgaagaataactcctcctcctccttaagaagagta 871

RESULT 12

AAK92006

ID AAK92006 standard; cDNA; 849 BP.

XX

AC AAK92006;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human cDNA 5'-end sequence, SEQ ID NO: 466.

XX

KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX

OS Homo sapiens.

XX

PN EP1130094-A2.

XX

PD 05-SEP-2001.

XX

PF 07-JUL-2000; 2000EP-0114089.

XX

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

XX

XX (HELI-) HELIX RES INST.

PA Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX

PT 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX
PS Claim 2; SEQ ID NO 466; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones, 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is the nucleotide
CC sequence of the 5'-end of a cDNA provided in the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;

Query Match 16.2%; Score 574.2; DB 22; Length 849;
Best Local Similarity 98.2%; Pred. No. 3.5e-146;
Matches 642; Conservative 0; Mismatches 6; Indels 6; Gaps 6;
Qy 2555 ttgggggaacatttctctacacacccatacttctctgcccagcttggagcagacagctat 2614
Db 198 ttgggggaacatttctctacacacccatacttctctgcccagcttggagcagacagctat 257
Qy 2615 cgctttacaacatttgaagcctactcactctctagaccaggaagtggatattgccaag 2674
Db 258 cgctttacaacatttgaagcctactcactctctagaccaggaagtggatattgccaag 317
Qy 2675 gtctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 2734
Db 318 gtctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 377
Qy 2735 tgcctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 2794
Db 378 tgcctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 437
Qy 2795 ttttacagatccagatgaccagctctcgaggttgcttctatgattaccacagagacctct 2854
Db 438 ttttacagatccagatgaccagctctcgaggttgcttctatgattaccacagagacctct 497
Qy 2855 acaatcacctggaggagacagatcgcccgccagcctctacgtcgccctggttctctca 2914
Db 498 acaatcacctggaggagacagatcgcccgccagcctctacgtcgccctggttctctca 557
Qy 2915 ccattgttgcctcagcttcccgctgggattcgtagcagagcttttgatgatttttc 2974
Db 558 ccattgttgcctcagcttcccgctgggattcgtagcagagcttttgatgatttttc 617
Qy 2975 ttcagggaacagaggtcatattaaagtggctttaaagctgttgggaagccataagccct 3034
Db 618 ttcagggaacagaggtcatattaaagtggctttaaagctgttgggaagccataagccct 677
Qy 3035 tgattctgcagcatgaanacattagaaccatagttgactttataaaaagcagcgtaccca 3094
Db 678 tgattctgcagcatgaanacattagaaccatagttgactttataaaaagcagcgt-ccca 736
Qy 3095 accttggcttggtagcagatgaaagaccatcaatcaatcaggtta-tttgaaatggacatc-gc 3152
Db 737 accttggcttgg-tagcagatgaaagaccatcaatcaatcangtatatttgnaatggacatcgg 795
Qy 3153 taacaggttacaagcttatg-aagttagttacca-cgtcccttcaagagaactt 3204
Db 796 taacaggttacaagcttatg-aagttagttacca-cgtcccttcaagagaactt 849

RESULT 13

AAK93666

ID AAK93666 standard; cDNA; 849 BP.

XX AAK93666;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human cDNA clone representative sequence, SEQ ID NO: 2126.
DE
XX
XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX EPI130094-A2.
PN
XX
XX 05-SEP-2001.
PD
XX
XX 07-JUL-2000; 2000EP-0114089.
PF
XX
XX 08-JUL-1999; 99JP-0194486.
PR
XX
XX 11-JAN-2000; 2000JP-0118774.
PR
XX
XX 02-MAY-2000; 2000JP-0183765.
PR
XX
XX (HELI-) HELIX RES INST.
PA
XX
XX Ota T. Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
PI
XX
XX WPI; 2001-524255/58.
DR
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
PT
XX
XX Example 11; SEQ ID NO 2126; 1380pp + sequence listing; English.
PS
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used as the
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;

Query Match 16.2%; Score 574.2; DB 22; Length 849;
Best Local Similarity 98.2%; Pred. No. 3.5e-146;
Matches 642; Conservative 0; Mismatches 6; Indels 6; Gaps 6;

Qy 2555 ttgggggaacatttctctacacacccatacttctctgcccagcttggagcagacagctat 2614
Db 198 ttgggggaacatttctctacacacccatacttctctgcccagcttggagcagacagctat 257
Qy 2615 cgctttacaacatttgaagcctactcactctctagaccaggaagtggatattgccaag 2674
Db 258 cgctttacaacatttgaagcctactcactctctagaccaggaagtggatattgccaag 317
Qy 2675 gtctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 2734
Db 318 gtctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 377
Qy 2735 tgcctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 2794
Db 378 tgcctcagctttgtagcagcatttggcttctctatagtgaggaagagcggttttaaaa 437
Qy 2795 ttttacagatccagatgaccagctctcgaggttgcttctatgattaccacagagacctct 2854
Db 438 ttttacagatccagatgaccagctctcgaggttgcttctatgattaccacagagacctct 497

QY 181 accaagcaagtcgctgttgcgtttccacccctgtgactgagatgtgaacccctgagccagg 240
|||||
Db 13013 accaagcaagtcgctgttgcgtttccacccctgtgactgagatgtgaacccctgagccagg 13072
|||||
QY 241 agaaqtcaacagtgagtcctcctgatctattccagcatctttgagtgcaagcctcagcgt 300
|||||
Db 13073 agaaqtcaacagtgagtcctcctgatctattccagcatctttgagtgcaagcctcagcgt 13132
|||||
QY 301 gtccacaactgattccacaagtcagtcacccaagtactttgtctgtgattaaggaa 360
|||||
Db 13133 gtccacaactgattccacaagtcagtcacccaagtactttgtctgtgattaaggaa 13192
|||||
QY 361 gacgtgtccaccgacagatgtctgtatgtgttcaaaagccgatgatacaaaaagt 420
|||||
Db 13193 gacgtgtccaccgacagatgtctgtatgtgttcaaaagccgatgatacaaaaagt 13252
|||||
QY 421 cctgagatcatcagtcctcctcgtcagcgagg 453
|||||
Db 13253 agtgagatggagatccaaaagactaagtgagg 13285
|||||

RESULT 15

ID AA215229 standard; cDNA; 696 BP.

AC AA215229;

DT 12-OCT-1999 (first entry)

DE Human gene expression product cDNA sequence SEQ ID NO:2698.

XX Human; gene; gene expression product; diagnosis; therapy; probe;

KW detection; mapping; tissue typing; profiling; forensic; cancer;

KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

XX Homo sapiens.

PN WO9938972-A2.

XX 05-AUG-1999.

XX 28-JAN-1999; 99WO-US01619.

PF 03-APR-1998; 98US-0080666.

PR 28-JAN-1998; 98US-0072910.

PR 24-FEB-1998; 98US-0075954.

PR 31-MAR-1998; 98US-0080114.

PR 03-APR-1998; 98US-0080515.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

XX Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
XX Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
XX Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
XX Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
XX Stache-Crain B, Sudduth-Klinger J, Williams LT;

DR WPI; 1999-494092/41.

XX Novel human genes and their expression products which are

PT differentially expressed in different cell types

XX Claim 1; Page 1311; 2479pp; English.

XX The present invention describes a library of human polynucleotides
XX comprising the sequences given in AA215232 to AA21779. Also described is
XX a method of detecting differentially expressed genes correlated with the
XX cancerous state of a mammalian cell, comprising detecting at least one
XX differentially expressed gene product in a test sample from a cell
XX suspected of being cancerous, where the gene product is encoded by one
XX of the 5248 polynucleotide sequences given in AA215232 to AA21779. The
XX polynucleotides can be used as a source of primers and probes, which can

CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.

XX Sequence 696 BP; 189 A; 157 C; 178 G; 142 T; 30 other;

Query Match 10.2%; Score 362; DB 20; Length 696;

Best Local Similarity 100.0%; Pred. No. 2.8e-88;

Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3194 aagaagaacttatcctctcctcctcagtgacacacaaagaatggataaattagaga 3253
|||||

Db 59 aagaagaacttatcctcctcctcagtgacacacaaagaatggataaattagaga 118
|||||

QY 3254 aaaccaacagcagcttacgcaaaacagacaccttgacctcttgacaggttgagcaa 3313
|||||

Db 119 aaaccaacagcagcttacgcaaaacagacaccttgacctcttgacaggttgagcaa 178
|||||

QY 3314 atggtagatccaaagccttgagccaccattgagaagctcctgagcagtgagcaagc 3373
|||||

Db 179 atggtagatccaaagccttgagccaccattgagaagctcctgagcagtgagcaagc 238
|||||

QY 3374 tgaagcagggccatgcttacccttagaactggagcgtcgccctgctgcagacggtggagg 3433
|||||

Db 239 tgaagcagggccatgcttacccttagaactggagcgtcgccctgctgcagacggtggagg 298
|||||

QY 3434 agctgcggcgaggcagcagaccagcagcgggagcctgagtcagcagcccgagc 3493
|||||

Db 299 agctgcggcgaggcagcagaccagcagcgggagcctgagtcagcagcccgagc 358
|||||

QY 3494 ccacggcgactgacagctctgcaggagagattgcaacacccatcccacactgccaggcc 3553
|||||

Db 359 ccacggcgactgacagctctgcaggagagattgcaacacccatcccacactgccaggcc 418
|||||

QY 3554 tt 3555

Db 419 tt 420

Search completed: September 9, 2002, 19:20:22

Job time: 17600 sec

GenCore version 4.5	
Copyright (c) 1993 - 2000 Compugen Ltd.	
OM nucleic - nucleic search, using sw model	
Run on:	September 9, 2002, 19:16:17 ; Search time 8477.27 Seconds (without alignments) 8775.683 Million cell updates/sec
Title:	US-09-762-311-4_COPY_176_3730
Perfect score:	3555
Sequence:	1 atgggaacaataacattcac.....tcccacactgtccaggcctt 3555
Scoring table:	IDENTITY_NUC Gapop 10.0 , Gapext 1.0
Searched:	1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters:	3595312
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries
Database :	GenEmbl.* 1: gb_ba.* 2: gb_hcg.* 3: gb_in.* 4: gb_om.* 5: gb_ov.* 6: gb_pat.* 7: gb_ph.* 8: gb_pl.* 9: gb_pr.* 10: gb_ro.* 11: gb_sts.* 12: gb_sy.* 13: gb_un.* 14: gb_vl.* 15: em_ba.* 16: em_fun.* 17: em_hum.* 18: em_in.* 19: em_nu.* 20: em_om.* 21: em_or.* 22: em_ov.* 23: em_pat.* 24: em_ph.* 25: em_pl.* 26: em_ro.* 27: em_sts.* 28: em_un.* 29: em_vi.* 30: em_htg_hum.* 31: em_htg_inv.* 32: em_htg_other.* 33: em_htgo_inv.*

ALIGNMENTS	
RESULT 1	BC014529 3023 bp mRNA linear PRI 19-DEC-2001
LOCUS	BC014529 Homo sapiens, clone IMAGE:3843156, mRNA.
DEFINITION	BC014529
ACCESSION	BC014529.1 GI:17939551
VERSION	1 (bases 1 to 3023)
KEYWORDS	human.
SOURCE	ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. Strausberg,R. Direct Submission Submitted (24-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cdna Library Preparation: Rubin Laboratory
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REMARK	
COMMENT	

SUMMARIES	
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	
Result NO.	Score Match Length DB ID Description

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 Contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 14 Row: c Column: 16
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Source

Location/Qualifiers

1..3023
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3843156"
 /tissue_type="placenta, choriocarcinoma"
 /clone_lib="NIH_MGC_21"
 /lab_host="pH10B-R"
 /note="Vector: pOTB7"

BASE COUNT 883 a 712 c 740 g 688 t

ORIGIN

Query Match 77.4%; Score 2751; DB 9; Length 3023;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 805 attgagaacacacctattagcgacacacatttgcagccacacagatatcgaggaaat 864
 DB 10 ATTGAGAACCCCTCATAGCGACACATATGTGCGCCACACACATATCGAGGAAAT 69

QY 865 cgaactatgctcttcacgattggccagctcgaagttaacctatcactcagtcctgacacaaa 924
 DB 70 CGAACTATGCTCTACGATTGGCCAGTCTGAAAGTTTACCTCATCAGTCTGACACAAA 129

QY 925 aaaaatgacattgagaaaaatttaaggagatatctttgctcctcagggcagcagacac 984
 DB 130 AAAATAGCATTTGGAGAAAATTTTAAAGGAGATATCTTTTTCCTCAGGCGATCAGACAC 189

QY 985 gtggaccactttgggtttatctgcgggagcttcgggaggtgcgggcttcaattttgtc 1044
 DB 190 GTGGACCACTTTGGGTTTATCTCTCGGGAGCTCTCCGGAGGTGCGGCTTTCATTATTGTC 249

QY 1045 tgttacggttttcagtgacaaaatgagctcgtggttgatgaaattatgatgacctgaaa 1104
 DB 250 TGTTAGCTGTTTCAGTGCACAAATGAGGCTCTGGTTGATGAAATTTATGATGACCCCTGAAA 309

QY 1105 caggcttcacggtggccagtcgacgacagcagctaaagcgccagccagctgtgtgag 1164
 DB 310 CAGGCTTTCAGGTGGCCGAGTGCAGACAGCAGCTAAGCGCCAGCCACGCTGTGTGAG 369

QY 1165 ggctgccccctgcaagcctgcacagctctgtgagagatagagggaattcttc 1224
 DB 370 GGCTGCCCCCTGCAAGCCCTGCACAGCTCTGTGAGAGGATAGAGGGAATTTCTTCC 429

QY 1225 aaaaacaaactagaactgcaaaagcaccctgacgacatttaaccaatcagagcaggcgact 1284
 DB 430 AAAACAAAATAGAACTGCAAAAGCACCCTGACGACATTTAACCAATCAGGAGCAGGCGACT 489

QY 1285 atttttgaagaggttcagaaattgagaccgagaaatgagcagcagagagaattgatt 1344
 DB 490 ATTTTGAAGAGGTTTCAAAAATTTGACACCGAGAAATTTGAGCGGAGAGATGAATTTGATT 549

QY 1345 attctttcttgatgttttatgtgaagaaacagaaacacatcccatattggggag 1404
 DB 550 ATTTCTTTCTGAGATGTTTATATGAGAGAAACACAGAAAGAACACATCCATATTTGGGAG 609

QY 1405 atgaacagacatcgcagatggcagcagagaataattgggaagtgaattacacccagtgcc 1464
 DB 610 ATGAAGCAGACATCGGAGATGGCAGCAGAGATAATTGGGAAGTGAATTTACCAACCCAGTGCC 669

QY 1465 actcgttttagctagatatgctgaaaaaacaagcaaaagatctttaacagagtttta 1524
 DB 670 ATCTGATTTAGCTAGATATGCTGAAAAACAAGCAAAAGAGATCTTTACAGAGTCTTTA 729

QY 1525 gaaagtatttttcccgggggtataaaagccagaggcctgcaggaacacactccatcagtgtg 1584
 DB 730 GAAAGTATTTTCTCCCGGGTAAATAAGCCAGAGGCTGCAGGAACACTCCATCAGTGTG 789

QY 1585 gatctgatatgctcctcgtctagttacattaaacacacagcaaaagagcactcgtgtgtg 1644
 DB 790 GATCTGATAGTCTCCCTGTCTAGTACATTAAGTAACACAGCAAAAGAGGCATCTGTGTGT 849

QY 1645 gaaaaggaggccttgcccatctctcagagagctcctttaaagctcctcctcgaggagac 1704
 DB 850 GAAAGGAGGCTTGGCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCCTCGGAGGAC 909

QY 1705 ctgtccagtgactcggagagtgatctctccagaaagcagctcctcgtctgcgccccagag 1764
 DB 910 CTGTCCAGTGACTCGGAGAGTCTATCTCCAGAAAGACAGCCAGCTCCGCTGTCCGCCAGCAG 969

QY 1765 gcttcagagaggcagcaaaacacctgagtcacttcccatcgaatgccaggaacactcca 1824
 DB 970 GCCTTCAGGAGCGAGCAAAACACCTTGAGTCACTTCCCATCGAATGCCAGGAACCTCCA 1029

QY 1825 caacctgcccgggggtccccgggggtttcgcaaaaggaaacttatgaggtatcactcagtg 1884
 DB 1030 CAACCTGCCCGGGGTCCCGGGGGTTTCGCAAAAGGAANAATTTATGAGGTATCACTCAGTG 1089

QY 1885 agcacagagacgcctcatgaacgaaaggactttgaaatccaaagcaaacactcttggatg 1944
 DB 1090 AGCACAGAGACGCTCATGAACGAAAGAGACTTTTGAATCCAAAGCAAAACCATCTTGTGTGAT 1149

QY 1945 tctgtgggaactcctgtgaagaccgagagcattcctcagcagcagacagatattcctccga 2004
 DB 1150 TCTGTGGGACTCCTGTGAAGACCCGGAGGCATCTCTGGAGGCGCAGATATTTCCTCCGA 1209

QY 2005 gtggaccaccccgagagcgctgcgattcttccagcagatatgaagattattcagagctg 2064
 DB 1210 GTAGCCACCCCGCAGAAAGCGTGCATTTCTCCAGCAGATATCAAGATTATTTCAGAGCTG 1269

QY 2065 ggagagcttccccacgatatctccttttagaacacagttgttgaaagatgggccccttggcccc 2124
 DB 1270 GGAGAGCTTCTCCCGACGATCTCCTTTAGAACCACTTGTGGAAGATGGGCTTTGGCCCC 1329

QY 2125 ccacagagaaaaaagagagacatcgtgagctccgagagctgtggaagaaaggctatt 2184
 DB 1330 CCACCAGAGAAAAGAAAAGGACATCTCTGAGCTCCGAGAGCTGTGGCAAAAGGCTATT 1389

QY 2185 cttcaacagatactgctctagaatggagaggaagaaatcagagctccaagcctctgaa 2244
 DB 1390 CTTCAACAGATACTGCTGTCTTAGAATGGAGAGGAAATCAGAAGCTCCAAGGCTCTGAA 1449

QY 2245 aatgatttctgaacaaagcgcctgaagctcgattatgaagaaataactcctctctttaa 2304
 DB 1450 AATGATTCTGTGAACAAAGCCCTGGAAGCTCGATTATGAAGAAATTTACTCCTCTCTTAAA 1509

QY 2305 gaagtaactcacagtgtggaaaaagatgcttagcactccaggaagatcaaaaaataagttt 2364
 DB 1510 GAAGTAACCTACAGTGTGGAAAAGACATGCTTAGCACTCCAGGAAGATCAAAAATTAAGTTT 1569

QY 2365 gacatgaaaaaatagcactcggctgtgtggcaagggtgtgccagctcatcaccaggtgaa 2424
 DB 1570 GACATGAAAAAATGCACTCGGCTGTGGCAAGGTGTGCACGTCTATCACCAGGTGAA 1629

QY 2425 atctgaaatttctagctgagcaattccaccttaaacaccagtttcccagcaaacagcag 2484
 DB 1630 ATCTGAAATTTTCTAGCTGAGCAATTTCCACTTTAAACACCAAGTTTCCCAACCAACAGCAG 1689

QY 2485 ccaaaagatgtccatcacaaaagaactctlaaagcagctgacttcccagcagcagtgatt 2544
 DB 1690 CCAAGGATGTGCCATACAAAGAACTCTTAAAGCACTGACTTCCCAAGCAGCATGCGATT 1749

Db 892 TTCAACCATGTGAGTGTGGTCCAGAACGGACTGGGAAGCCGCCACCGGGCAGCCATCA 951
Qy 658 ggcgccacagggagccagagcctgtgscagagcccatgcgaagtcccttcccagccc 717
Db 952 GC-----GCTGGCCCGCAGGCCCATCGCAAAATCTTCTCACAGCCT 993
Qy 718 ggcctgcctgcctgccttttaggaagagcgtgcaggatgggggctccgaagcagcggc 777
Db 994 GGACTGCGCTCGCTGGCCCTTCAAGGAAGAGATTCCAGGACGCTAGCCTCGCAGTAGCAC 1053
Qy 778 ttcttcagctccttcagagagagacattgagaaccacctcattagcggacacaattatt 837
Db 1054 TT-----TAGTCTCCTTTGACAAATGACATAGAGAACCACTCATCGTGGGCACAATG 1107
Qy 838 gtgcagccacagatatcgaggaataatgaactatgctcttcacgatgtgccagctgaa 897
Db 1108 GTTCAGCCCAACAGATGAGAGGAAACCGAATATGCTGTTTCCAGATGGCCCATCTGNA 1167
Qy 898 gttacctcatcagtcctgcaccccaaaaaatagcattggagaaaaattttaaggagata 957
Db 1168 GTTTACCTCATCAGTCTGACACCAAAAAGATTGCACGTGGAGAAAAATTTTAAGGAGATA 1227
Qy 958 tccctttgctcagggcatcagcaacgtggaccacctttgggtttatctgtcgggagctct 1017
Db 1228 TCCTTTTGCTCTCAGGGCATCAGACATGTGGACCACTTTTGGATTTCATCTGCCGAGAGTGC 1287
Qy 1018 tc-----cggagggtggcgccttcaatttctgtctgctgtttcagtgcaacaaatgag 1071
Db 1288 TCGGTGGCGGCGAGTGGCGGCTTTCAATTTTGTCTGTTACTGTTCAGTGCACAAATGAA 1347
Qy 1072 gctctggttgataaattatgatgacctgaacggccttcacggtggcgcagtgcaag 1131
Db 1348 GCTCTGTTGACGAGATCATGTGACTCTGAAGCAGGCTTTTACGGTAGCTCGGTGTCAG 1407
Qy 1132 cagaaagctaaaggccagcctcagctgtgtgagggtgccccctgcacaaagcctgcacaa 1191
Db 1408 CAGACGGCTAAGGCACACGCCAGCTCTGTGAGGGGTGCCCTTGCAGGGCTGCAACAAG 1467
Qy 1192 ctctgtgagagatagaggaatgaattcttccaaacaaactagaactgcaaaagcac 1251
Db 1468 CTTCTGGAAGGATAGAGGGAATGAATTCATCTAANAACCAATTAGAACTCCAGAGCAC 1527
Qy 1252 ctgacgacattcaacatcagagcagcaggcagctatttttgaagaggttcagaaattgaga 1311
Db 1528 TTCAACCACTACCAATCAGGAGCAGGCCACCATATTCGAGAGGTTTTCAGAAATTTGAGA 1587
Qy 1312 ccgagaaatgagcagcagagagaatgaattatttcttcttcagagatttatatgaa 1371
Db 1588 CCAAGAAACGAGCAGCGAGAGATGAATTAATTTATTTCTTTCTGAGGTGCTTATATGAA 1647
Qy 1372 gagaaacagaaagaaacacatccatattggggagatgaaacagacatcgcagatggcagca 1431
Db 1648 GAGAGCAAAAAGCAGCACAGCCACACTGGGGGCCCAAGCAGACACTACAGGTGGCAGCA 1707
Qy 1432 gagaaattggaagtgaattaccaccagctgcacactcatttaggtagatatgtctgaaa 1491
Db 1708 GAGAAATTTGGGAGTGACCTGCCACAGCTGTAGCGGTTTCAGGTTAGATTGCTGTAAG 1767
Qy 1492 acaagacaaagagatctttaacagagcttttagaagatttttcccggggtaataaa 1551
Db 1768 AACAGACAAAGAGGTCTTTAACAGAGTCCCTAGAGAGCAATTCGTGCCGGGGTAATATA 1827
Qy 1552 gccagagggcctgcaggaaacactcactcagtggtgatcttgatagatccctctgtctagaca 1611
Db 1828 GCCAGAGGCTGCAGGACCAATTCGCCGCGAGTGTGGATCTGGACAGCTCCACTTCTAGTACT 1887
Qy 1612 ttaagtaacaccagcaagagcactctgtgtgtgtaaaaggagccttgcacctctctgag 1671
Db 1888 CTAAGTAACACCAAGAGAGCTGTCCATGGGTGACAAGAGGCTTCCCGCTCTCTGAG 1947
Qy 1672 agtctcttaagctcctcgtcctcaggaagacctgtccagtgactcgaagatcatctc 1731
Db 1948 ACCTCTCTCAAGCTCTTGCTCTCAGATGATGACCTGTGCCAGTGACTCAGAGGCCACATTT 2007

Qy 1732 ccagaagagccagctcccgctgtcggccagcaggcgttcaggaagcagcaaacccctg 1791
Db 2008 GCAGAAAGAGTCTGCCCTGTGTTGTCAACCCAGCAGGCGTTTCAGAAGGAGAGCAACCCCTG 2067
Qy 1792 agtacttcccactcgaatgccaggaaacctccaaacctgcgggggggttccccgggggtt 1851
Db 2068 AGTCATTTTCCCAAGTAGAGTGGCCCTCGCCTCCAGAACCTTCCAGAGCTCTCCAGGGGTC 2127
Qy 1852 tcgcaaaagaaacttatgagggtatcactcagtgagcacagacagcctcctcatgaagcaag 1911
Db 2128 TCTCAAAGAAACTCATGCGGTACCACTCCGTGAGCACAGAGACGCTCATGACGCAAG 2187
Qy 1912 gactttgaatccaaagcaaacatcttgtgtattctgtgtgggactcctctggaagcccg 1971
Db 2188 GACTTTTGAATCCAAAGCAAAACCACTGGGTGACACAGATGGGACCCCTGTAAGACCCGG 2247
Qy 1972 aggtcttctcgtgagcagcagatatctcctccagtagtgcacccccccagcagaagcgtagat 2031
Db 2248 CGGCACCTCGTGGAGACAGCAGATATTCCTTCGAGTGGCCACTCCACAGAAAGCTTGTGTAC 2307
Qy 2032 tcttcagcagatatgaagattattcagagctgggagagcttccccccagctctccttta 2091
Db 2308 TCCCGAGCAGATATGAGATTTATTCGAGCTGGGAGAGCTCCCTCCAGCTCCCTTTTA 2367
Qy 2092 gaaccagtttgtgaagatggggccctttggccccccacccagagaaaagagagacatct 2151
Db 2368 GAACCGGTGTGTGAGACGCGCCCATTTGG-CCAGTACAGGAAGAAAGAGAGAGCTCA 2426
Qy 2152 cgtgagctcggagagcgtgtggcaaaaggctatttctcaacagatactg-ctgcttgaagt 2210
Db 2427 CCGGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTGCAGCAGATCTCGCTCGTAGGAT 2486
Qy 2211 ggagaagaaaaatcagaagctccaagcctctgaaaatgattgttgaaacaaagcgccgtgaa 2270
Db 2487 GAGAGAGGAGATCAGAGCTACAGCCTCTGAAAACGATTTGCTGAAACAACGCGCTCAA 2546
Qy 2271 gctgattatgaagaaattactccctgtctaaagaagtaactacagatgtgtgggaaagat 2330
Db 2547 GCTTGACTATGAAGAAATCAGTCCCTGTCTTAAAGAAGTCACTACAGTGTGGGAAAGAT 2606
Qy 2331 gcttagcactccagagagatcaaaatgaattgacatgggaaaaaatcactcggtgt 2390
Db 2607 GCTTAGACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAGTGCATCAGCTCT 2666
Qy 2391 tgggcaaggtgtgccacgtcatccacgaggtgaaaatcttgaaaattctagctgagcaatt 2450
Db 2667 TGGCAAGCTGTGCCAGCTCATCACCGAGGTGAGATCTGGAATTTCTAGCTGAGCAGT 2726
Qy 2451 ccaccttaaacaccagtttccagcaaaacagcagcaaaaggatgtgccaatacaaaagact 2510
Db 2727 CCACCTTAAACCCCATTTCTTAGTAAACAGCAGCCAAAGGAGCTGCCCTTACAAAGCT 2786
Qy 2511 cttaagcagctgaactcccagcagcatgctatttattgacctggggcgaaccttcc 2570
Db 2787 CTTGAAGAAGATGACCTCGCAGCAGCAGCCCATCTCTCATCGACCTCGGGGGAACCTTCC 2846
Qy 2571 tacacacccatacttctctgcccagcttggagcagagcagctatcgcttttaaacatttt 2630
Db 2847 AACACATCCATACTCTCTGCCCAGCTTGGAGCAGGTGAGTGTCTACTTTTACAACTTCT 2906
Qy 2631 gaaggcctactcacttcttagaccaggaagtgggatatgtccaaggtctcagcttttagc 2690
Db 2907 GAAGGCTACTCGCTTCTGACAGAGGTTGGATTAATGCTCCAAAGGTCCTAGCTTTTGGC 2966
Qy 2691 aggcatttgccttctcatatgagtgaggaagagcggtttaaaatgtcgaatttctgat 2750
Db 2967 AGGCATTTTCTTCTTACATGAGTGGAGAGAGGCTTCAAGATGCTCAAGTTCCTGAT 3026
Qy 2751 gtttgacatggggtcgggaacagtatcgggccagacatgatatttttacagagatccagat 2810
Db 3027 GTTTGACATGGGCTCGGGAAACAGTATCGGCCAGACATGATTATTTTGCAGATCCAGAT 3086

Db 421 TCCTGTGTGAAAGAGGCGCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCCTCGGCTCC 480
QY 1696 tcgaggagactgtccagtgactcggagagatcatctcccaagaagagccagctccgctgtcg 1755
Db 481 TCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAAGAAGAGCAGCTCGCTGTGCG 540
QY 1756 cccacagagccttcagagagcgagcaaacacccctagtcacttcccacatgaatccag 1815
Db 541 CCCCACAGGCGCTTCAGGAGCGGAGCAACACCCCTGAGTCATCTCCCAATCGAATGCCAG 600
QY 1816 gaacctcccaaacctgccggggggtcccggggggtttcccaaggaaacttatgaggtat 1875
Db 601 GAACCTCCCAACACTGCCCGGGGTCCCGGGGGTTTCGAAAGGAAACTTATGAGGTAT 660
QY 1876 cactcagtgagcagagagagcctcatgaacgaagaactttgaatccaaagcaaacct 1935
Db 661 CACTCAGTAGCAGACAGAGCGCTCATGAACGAAAGAGACTTTGAATCCAAAGCAAACTAT 720
QY 1936 ctgggtgactctgtgggactcctgtgaagaccgagagcattcctggagcagcagata 1995
Db 721 CTTGGTGATCTCGTGGGACTCTGTGAAGACCCGGAGGCATTCCTTGGAGGAGCAGATA 780
QY 1996 ttccctcagtagtagccaccgcgagaaggcgtgcgattcttccagcagatatgaagattat 2055
Db 781 TTCTCTCGAGTAGCACCCCGCAGAGGCGTGGGATTCCTCCAGCAGATATGAAGATTAT 840
QY 2056 tcagagctggagagcttcccaccagcatctctttagaaccagtttgtgaagatgggccc 2115
Db 841 TCAGAGCTGGGAGAGCTTCCCCACCATCTCTTTTGAACACAGTTTGTGAAGATGGGCC 900
QY 2116 ttggccccccacagagaaagaaagacatctcgtgagctccgagagctgtggcaa 2175
Db 901 TTTGGCCCCCACCAGAGAAAGAAAGGACATCTCGTGAGCTCCGAGAGCTGTGGCAA 960
QY 2176 aaggctattcttcaacagatactgtgcttagaatggagaaggaaatcagaagctccaa 2235
Db 961 AAGGCTATTCTTCAACAGATACCTGCTGTAGAAATGGAGAAGGAAATCAGAAGCTCCAA 1020
QY 2236 gcctctgaaatgatitgtgaacaagcgcctgaactcgattatgaagaataactccc 2295
Db 1021 GCCTCTGAAATGATTGCTGAACAAGCGCTGAAGCTCGATTATGAAGAAATTACTCCC 1080
QY 2296 tgtctlaaagaagtaactacagtggtgggaaagatgcttagcactccaggaagatacaaa 2355
Db 1081 TGTCTTAAAGAACTACTACAGTGTGGGAAAGATGCTTAGCAGCTCCAGGAAGATCAAA 1140
QY 2356 attaatgtgacatgaaataatgcactcggctgtgtggcaaggtgtgcacagtcacac 2415
Db 1141 ATTAAGTTTGACATGGAAAAAATGCACTCGGCTGTGGGCAAGGTGTGCCACGTCATCAC 1200
QY 2416 ctaggtgaaatctggaaattctagctgagcaattccacttaaacaccagttcccagc 2475
Db 1201 CGAGGTGAAATCTGGAATTTCTAGCTGACCAATTCACCTTAAACACCAATTTCCAGC 1260
QY 2476 aaacagcagccaaagtagtgccatatacaagaactcttaaacagcgtgacttcccagcag 2535
Db 1261 AAACAGCAGCAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCAGCAG 1320
QY 2536 catgcattcttattgacttggggaacacttccctacacaccaacttctctgtgccag 2595
Db 1321 CATGCGATTCTTATTGACCTTGGCGAACTTCTTCCACACACCCACTTCTCTGCGCCAG 1380
QY 2596 ctggagcagcagcagctatcgtttcaacattttgaagcctactcactcttagaccag 2655
Db 1381 CTTGGAGCAGGACAGCTATCGCTTTCAACATTTTGAAGCCCTACTACTCTTAGACCAG 1440
QY 2656 gaagtgggattgccaagctcagcttggtagcaggcaatttgccttctcataggt 2715
Db 1441 GAAGTGGATATGCCAAGCTCTCAGCTTTGTAGCAGGCAATTTGCTTCTCATATGAGT 1500
QY 2716 gaggaaagcggttttaaaatgcctcaagttctctgatgtttgacatggggctgcggaaacag 2775
Db 1501 GAGGAAGAGCGCTTAAATGCTCAAGTTCTTATGTGTTGACATGGGGCTGCGGAACACAG 1560

QY 2776 tatcgccagacatgattatttaccagatccagatgaccagctctcagggtgtctcat 2835
Db 1561 TATCGGCCAGACATGATTATTACAGATCCAGATACAGCTCTCGAGGTTGCTTCAT 1620
QY 2836 gattaccacagagacacctacacatcacctcgggagcagcagatcgccccagcctctac 2895
Db 1621 GATTACCACAGAGACCTCTACAATCACCTTGAGGAGCAGAGATCGCCCCAGCCTCTAC 1680
QY 2896 gctgcctcctgttctccaccatgttgctcacagctcccgctgggattctgtagccaga 2955
Db 1681 GCTGCCCTGTGTTCTCACCATGTTTGCTTACAGTTCCCGTGGGATTCGTTAGCCAGA 1740
QY 2956 gctcttgatatatttcttcagggaacagaggtcatattttaaagtgccttcaagtctg 3015
Db 1741 GCTTTGATATGATTCTTTCAGGAACAGAGGTCATATTAAAGTGGCTTTAAGTCTG 1800
QY 3016 ttgggaagccacaagaagccttgcagcatgaaacctagaacccatagttgacttt 3075
Db 1801 TTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAACCTAGAAACCATAGTTGACTTT 1860
QY 3076 ataaaagcagcctaccacccacttggcttggtagagatggaaaaagacatcaatcaggta 3135
Db 1861 ATAAAAGCAGCGTACCCCAACCTTGGCTTGGTACAGATGGAAGAACCATCAATCAGGTA 1920
QY 3136 ttggaatgagcatcgctaaacagttacaagcttatgaagttgagtcaccagctccctcaa 3195
Db 1921 TTTGAATGGACATCGCTAAACAGTTTACAGCTTATGAAGTTGAGTACCACGCTCTCAA 1980
QY 3196 gaagaacttatcgattcctcctcctcagtgacacacccaaagaaatggataaataagagaaa 3255
Db 1981 GAAGAACTTATCGATTCTCTCTCTCAGTGACCAACCAAGAAATGGATAAATTAGAGAAA 2040
QY 3256 accaacagcagcttaacgcaaacacacacttgacctccttgaaacagttgaggtggcaaat 3315
Db 2041 ACCAACAGCAGCTTACGCAACAGAACTTGACCTCTTGAACAGTTGGAGGTGGCAAT 2100
QY 3316 gctagatccaaagccttgagccaccattgagaagctcctgagcagtgagagcaagctg 3375
Db 2101 GGTAGGATCCAAAGCCTTGAGCCACCATTTGAGAACTCTTGAGCAGTGAGACCAAGCTG 2160
QY 3376 aagcaggccatgcttacccttagaactggagcgtcgccctcgtcagacagcgtggaggag 3435
Db 2161 AAGCAGGCCATGCTTACCTTAGAACTGGAGCGCTCGCCCTGCTGCAGACGGTGGAGGAG 2220
QY 3436 ctgcggcgagcagcagcagccagccagccgagccctgagtgacgcagcagcccgagccc 3495
Db 2221 CTGCGCGGGAGCGAGCGAGCCCGAGCCGAGCCCTGAGTGCACGACGCGCCGAGCCC 2280
QY 3496 acggcgactgacagctcgtcaggagagattgcaacacacatcccacactgtccagcctt 3555
Db 2281 ACGGGGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCACACTGTCCAGGCTT 2340

RESULT 5

AK027355
LOCUS Homo sapiens cDNA FLJ14449 fis, clone HEMBB1001684.
DEFINITION AK027355
ACCESSION AK027355
VERSION AK027355.1 GI:14041975
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly body cDNA to mRNA, clone lib:HEMBB1 clone:HEMBB1001684.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (sites)
Isoqai,T., Ota,T., Hayashi,K., Sudiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Makamatsu,A.,
Ishii,S., Kawai,Y., Salto,K., Yamamoto,J., Wakamatsu,A.,

Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2362)
 Isogai, T. and Otsuki, T.
 Direct Submission
 Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing,
 Research Association for Biotechnology; cDNA library construction,
 5'- & 3'-end one pass sequencing and clone selection; Helix
 Research Institute (supported by Japan Key Technology Center etc.)
 and Department of virology, Institute of Medical Science,
 University of Tokyo

FEATURES source

1. .2362
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEMB100168"
 /tissue_type="whole embryo, mainly body"
 /clone_lib="HEMB1"
 /dev_stage="embryo, 10 weeks"
 /note="cloning vector: pME18SFL3"
 40..2064
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB5057.1"
 /db_xref="GI:14041976"
 /translation="MLNKAARSLTLESILSRGNKARGLOPHSIVDLSDLSSTL
 SNPSKPSVCEKALPISESSFKLIGSSDSESHLPEEPAPLSPQOAFRRRANT
 LSHPFECPPQAPGSPGVSRKLMRYHSVETEPHERKDFESKRNHLDGSGGPV
 KRRHSWRQDI FIVRTAPQKACDSSSYEDYSELGPPRSPLEPCVDEGPFPPPEE
 KRTSLRELPGKAIQIILLMEKNOKLAQSENDLNKLKLDYEITPLKVEY
 TTVMKLSLPGRSKTKFDMKMSHAGVGVPRHRRGEIWKFLAEQFHLKHQPSKQO
 PKDVPYKELLQLTSSOHA TLIDLGRFTFPHYFSAQLGAGQLSLNKLKAYSLDQE
 VGYCOGLSFVAGILLHLMSEEAFAKMLKFLFMDGLRKYRPMIILIQMYOLSRLL
 HDYRDLNHEHEGEGSLYAPWELTFMFAQFPLGVFARVDFMFLQGETVIFKVA
 LSLGSHKPLILQHENLEITVDVFKSTLPNLGVMEKTIQNVFMDIAKQLQAYEVE
 YHVLQELIDLSPLSNORMDKLEKTNSSRLKQNLDLLEQLQVANGRIQSLQATIEKL
 LSSSEKQLQAMLTLELRSALLQTVELRRRSRPSDREPECTQPEPTGD"

CDS

657 a 581 c 578 g 546 t
 BASE COUNT
 ORIGIN

Query Match 59.4%; Score 2112; DB 9; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1444. agtgaattaccaccagtcgactcgatttaggtagatgctgaaacaagaagcaag 1503
 Db 1 AGTGAATTACCACCAGTCGACATCGATTTTAGGCTAGATATGCTGAAACAAGCAAG 60
 Qy 1504 agatcttaacagagctcttagaagattttgtccggggtataaagccagagcctg 1563
 Db 61 AGATCTTTACAGAGCTCTTTAGAAGATATTGTGTCGGGGGTATTAAGCCAGAGGCTG 120
 Qy 1564 caggaacactccatcagtggtgatctgtagctccctgtctgtagtacatgaagtaaacacc 1623
 Db 121 CAGGAACACTCCATCAGTGTGGATCTGGATAGTCCCTGTCTAGTACATTAAGTAACACC 180
 Qy 1624 agcaagagccatctgtgtgcaaaagagccttgccatctctgtgagagctctttaag 1683
 Db 181 AGCAAGAGCCATCTGTGTGTGAAAGAGGAGGCCTTGCCATCTGAGAGCTCTTTAAG 240
 Qy 1684 ctctcggtcctcgagagccttcagtgactcgagagctcatctccagagagacca 1743
 Db 241 CTCCTCGGCTCCTCGAGGACCTGTCCAGTACTCGGAGAGTCACTCCCAAGAGGCCA 300
 Qy 1744 gctccgctgtgcccagcagccttcagggagcggagcaaacaccctgagtcacttcccc 1803
 Db 301 GCTCCGCTGTGCCCCACAGGCCTTCAGGAGCGGAGCAACACCCTGAGTCACTTCCCC 360

Qy 1804 atcgaatgccaggaaacotccacaacctgcccgggggtccccgggggttttcgcaaaagaaa 1863
 Db 361 ATCGAATGCCAGGAACCTCCACAACCTGCCCGGGGGTCCCGGGGGTTTCGCAAAAGGAAA 420
 Qy 1864 cttatgagggtatacactcagtgagcacagagacgctcatgaacaaagagactttgaatcc 1923
 Db 421 CTTATGAGGTATCACTCAGTGAGCACAGACGCGCTCATGAACGAAGGACTTTGAATCC 480
 Qy 1924 aagcaaacaccatcttggtgattcttggtggaactcctctgaagaccggagagcattctctg 1983
 Db 481 AAGCAAAACCATCTTGTTGATCTCTGTGGGACTCTCTGTGAAGACCCGGAGGCATTCCTGG 540
 Qy 1984 aggcagcagatattctccagtagtagcccccgcagagggcgtgcgattcttccagcaga 2043
 Db 541 AGGCAGCAGATATTCTCCGAGTAGCCACCCCGCAGAGGCGTGGGATCTTTCAGGAGA 600
 Qy 2044 tatgaagattattcagagctgggagagcttccccccagatctctcttagaacaccgtttgt 2103
 Db 601 TATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCCAGCATCTCTCTTAGAACCAAGTTGT 660
 Qy 2104 gaagatgggccccttggtggccccccaccagagagaaagaaagacatctctgagctcoga 2163
 Db 661 GAAGATGGGCCCTTTGGCCCCCCCCACCAGAGAGGAAAGAAAGACATCTCGTGAGCTCGGA 720
 Qy 2164 gagctgtggcaaaaggctatttctcaacagatactgctgcttagaaggagagaaat 2223
 Db 721 GAGCTGTGGCAAAAGGCTATTCTTCAACAGATACTGCTGTCTTAGAATGGAGAGGAAAT 780
 Qy 2224 cagaagctccaaagcctctgaaaaatgattgctgacaagcgcctgaagctcgattatgaa 2283
 Db 781 CAGAAGCTCCAAGCCTCGAAAATGATTGCTGAACAAGCGCTGAAGCTGATTATGAA 840
 Qy 2284 gaaattactccctgtcttaagaaagtaactacagctgtgggaaaaagatgcttagcactoca 2343
 Db 841 GAAATTACTCCTGCTCTTAAAGAAAGTAACACAGTGTGGGAAAAAGATGCTTAGACACTCCA 900
 Qy 2344 ggaagatcaaaaaatgaatttgacatggaaaaaataagcactcggctgttgggcagaggtg 2403
 Db 901 GGAAGATCAAAAATTAAGTTTGACATGGAAAAATGACCTCGGCTGTGTGGCAAGGTGTG 960
 Qy 2404 ccagctcatcacccaggtgaaatctgaaatttctagctgagcaattccacacttaaacac 2463
 Db 961 CCAGCTCATCACCGAGGTGAATCTGGAATTTCTAGCTGAGCAATTTCCACCTTTAAACAC 1020
 Qy 2464 cagttccccagcaaacagcagcgaagatgctgccaatacaagaactctttaaagcagctg 2523
 Db 1021 CAGTTTCCCAGCAAAACAGCAGCCAAAGGATGTGCCATACAAGAACTCTTTAAAGCAGCTG 1080
 Qy 2524 actccagcagcagatgagattcttattgaccttggggaaacctttccacacccatcac 2583
 Db 1081 ACTTCCCAGCAGCATGCGATTCTTTATTGACCTTGGGGGAACTTTCTTACACACCCATAC 1140
 Qy 2584 tctctgccagcgttgagcagcagcagctatcgctttacacattttgaagcgctactca 2643
 Db 1141 TTCTCTGCCAGCTTGAGCAGGAGCAGCTATCGCTTTTACAACATTTTGAAGGCGCTACTCA 1200
 Qy 2644 ctctagaccaggaagtgggatatgccaaggtctcagctgttgtagcaggaatttgcctt 2703
 Db 1201 CTTCTAGACAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCAATTTTGCCT 1260
 Qy 2704 ctccatgagtgaggaagagggcgtttaaatgctcaagttctgattgttgacatgggg 2763
 Db 1361 CTTTATGATGAGTGAGGAAGAGGCGCTTTAAATGCTCAAGTTTCTGATGTTTGCATGATGGG 1320
 Qy 2764 ctgcggaaaacagtagtcggccagacatgattattttacagatccagatgtaccagctctcg 2823
 Db 1321 CTGCGGAAACAGATATCGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCG 1380
 Qy 2824 aggttgcttcattgattaccacagagacctctacaatacacctggggagcagagatcggc 2883
 Db 1381 AGGTTGCTTATGATTATCCACAGAGACCTCTACAAATCACCTGAGGAGGACGAGATCGGC 1440

Db 61 TCCCTTTAGACCGGTGTGAGGACGCGCCATTGTGGCCAGTACAGGAAGAAAGAGG 120
QY 2143 aggaatctctgtagctcgagctgtgtgcaaaagcttatttcttaacagatactgtctg 2202
Db 121 AAGACGTCTACGGGAGCTTCGAGAGCTGTGGAAGAGGCGCATCTTGACGAGATCTCTGTG 180
QY 2203 cttagaatsgagaagaaatcagaagctccaaagctctgaaatgatttctgaaacaag 2262
Db 181 CTCAGGATGGAGAGAGAAATCAGAAGCTACAAAGCCCTCGAAGACGATTTGCTGAACAAA 240
QY 2263 cgcctgaagctgattatgaagaataactccctgtcttaagaagtaactacagtggtg 2322
Db 241 CGCCTCAAGCTTGACATGATGAATCACTCCGTGTCTTAAAGAACTCACTACAGTGTGG 300
QY 2323 gaaaagatcttagactccaggaagatcaaaaataatttgaatggaataaaatgcac 2382
Db 301 GAAAGATGCTTAGACACTCCAGGAAGATCAAAATTAAGTTTGACATGGAAGAAAGTGCAC 360
QY 2383 tcggctgttgggcaaggtgtgcacgtcatccagaggtgaaatctgaaatttctagct 2442
Db 361 TCAGCTGTGGGCAAGGTGTGCCACGTCAACCGAGGTGAGATCTGGAATTTCTAGCT 420
QY 2443 gagcaattccacattaaacaccagatttcccagcaaacagcagccaaagatgtgcccatac 2502
Db 421 GAGCAGTTCCACCTTAACACCCATTCTCTAGTAACAGCAGCCAAAGCAGCTGCCTTAC 480
QY 2503 aaagaactcttaagacagctgacttcccagcagatgcgatttattttagcttggcgga 2562
Db 481 AAGAGCTCCTGAAGAAGCTGACCTCGCAGCAGCAGCCATTCTCATCGACCTCGGGCGA 540
QY 2563 accttccctacacaccatacttctctccagcttggagcagcagcagctatcgctttac 2622
Db 541 ACCTTTCCAAACATCCATCTCTCTGCCAGCTTGGAGCAGGTGACGTGTCACATTTAC 600
QY 2623 aacatttgaagcctactcaactcttagaccaggaagtggtgatatggcaaggtctcagc 2682
Db 601 AACATTTGAAGCCCTACTCGCTTCTGGACAGGAGGTTGGATCTGCCAAGGTCTCAGC 660
QY 2683 ttgtgagcaggaatttcttctctatagtgaggaagagcgttttaaaatgctcaag 2742
Db 661 TTTGTGGCAGGCATTTTGTCTCTCATAGTGAAGGAAGAGCGCTTCAAGATGCTCAAG 720
QY 2743 ttctgagttgacatgggctgcggaacagatcgcgccagacatgattattttacag 2802
Db 721 TTCTCTGATGTTGACATGAGGGCTCCGGAACACATATCGCCAGACATGATATTTGCGAG 780
QY 2803 atccagatgtaccagctctcgaggttgcttcatgattaccacagagacacctcaaatcac 2862
Db 781 ATCCAGATGTACCAGCTGTACGGCTCTCCACGATTTACCACCGAGACCTCTACAACCCAC 840
QY 2863 ctggagagcagagatcgcccccagcctctacgtgcgcccttggttctccaccatgttt 2922
Db 841 CTGGAAGACGACGAGATCGGCCCGCCAGCCTCTAGCGGGCTCCCTGGTTCTCACCGGTTC 900
QY 2923 gcttcacagttcccgctggagctcgtagcagagcttttgatatgatatttcttcaggga 2982
Db 901 GCCTCACAGTTCCTCCATCGGCTTTGTAGCCAGAGCTCTTTGATATGATCTCTTCAGGGA 960
QY 2983 acgaggtcatatttaagtggttttaagctctgttgggaagccataagcccttgattctg 3042
Db 961 TCAGAGTGCATATTTAAAGTAGCTTTTAAAGTCTTTTGGGAGCCATAAAGCCCTTGATTTCA 1020
QY 3043 cagcatgaaacctagaaccatagttgactttataaaagcagcgtaccacaccttgc 3102
Db 1021 CAGCATGAGAACCTGGAACCATCTGGTGGACTTCATAAAGAACACACTCCCCCAACCTGGGC 1080
QY 3103 ttggtcagatggaagaccatcaatcaggtatttgaatgacatcgcttaaacagtta 3162
Db 1081 CTGGTCAGATGGAGAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAGCAGCTC 1140
QY 3163 caagcttatgaagttgagttaccagctctcttcaagaagaacttatcgattctctctctc 3222

Db 1141 CAGCCTATGAGGTGAGTACACAGTGTCTCCAGAGGAGCTTATTAGTCTCGCCTCTC 1200
QY 3223 agtgcacaaaagaatggataataatagagaaaacaaacagcagcttaagcaaaacagaac 3282
Db 1201 AGTGAACAAACAAAGAAATGGAGAAATGGAGAAACCAACAGCAGCTTGGCCAAACAGAAC 1260
QY 3283 cttgacctcttgaacagttgaggtggcaaatggttagatccaaagccttggagccacc 3342
Db 1261 CTTGACCTCTCGAGCAGTGTGAGGTGGCAATGCTAGATCCAAAGCCCTTGAGCCACG 1320
QY 3343 attgaagaagctcctgagcagtgagcagcagctgaagcagcagcccatcttacccttagaactg 3402
Db 1321 GTAGAGAAACTTCTTACCAGCGAGAGTAAAGCTGAAGCAGGCTGCCGTGACCTGGAGGTG 1380
QY 3403 gagcgtggtgcccctgctgcagcagcgtgagagcgtcggcggtggagcagagccacc 3462
Db 1381 GAGCCTCGGCCCTCTGTCAGATGTTGGAGGAGCTGCGAGGAGCAAGCCCGGCCCGCCAGC 1440
QY 3463 gaccgggagcctgagtgcaagcagcccgagcccgagcggcgactgacagct 3512
Db 1441 ACTCCAGAGCCAGAGTGCACCCAGCTGGAGCCACAGGCGGATTGACCGCT 1490

RESULT 8
LOCUS AK057182 1788 bp mRNA linear PRI 31-OCT-2001
DEFINITION Homo sapiens cDNA FLJ32620 fis, clone STOMA2000386, highly similar to Bos taurus mRNA for linceln.
ACCESSION AK057182
VERSION AK057182.1 GI:16552778
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens stomach cDNA to mRNA, clone_lib:STOMA2
clone:STOMA2000386.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (sites)
AUTHORS Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagaatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEBO human cDNA sequencing project
Unpublished
TITLE 2 (bases 1 to 1788)
JOURNAL Isogai,T., Otsuki,T. and Sugiyama,T.
REFERENCE Direct Submission
AUTHORS Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
TITLE NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES
source Location/Qualifiers
1..1788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STOMA2000386"
/tissue_type="stomach"
/clone_lib="STOMA2"
/note="cloning vector: pME18SFL3"

BASE COUNT ORIGIN	481 a	412 c	434 g	461 t
Query Match Best Local Similarity Matches 999; Conservative	28.1% 99.9% 999;	Score 998.4; Pred. No. 1.3e-207; 0; Mismatches	DB 9; 1; Indels	Length 1788; 0; Gaps
Qy	2556	tggcggaacatttctctacacaccacttctctccagcttgagcagagacgctatc	2615	
Db	553	TGGCGGAACCTTCTCTACACACCATACTCTCTGCCAGCTTGAGCAGCAGCAGCTATC	612	
Qy	2616	gctttacaacatttgaagccttactcactcttagacagagaaagtggatattgccaagg	2675	
Db	613	GCTTTACAACATTTTGAAGGCTTACTCACCTCTAGACAGGAAGTGGATATGCCAAGG	672	
Qy	2676	tctcagctttgttagcagcatttgccttcttcatatgagtgaggaagcgctttaaatt	2735	
Db	673	TCTCAGCTTTGTAGCAGGCATTTTGTCTTCTCATATCAGTGAGGAAGCGGTTTAAAT	732	
Qy	2736	gctcaagtcttgatgtttgacatgggctgcggaacacagtatatggccagacatgattat	2795	
Db	733	GCTCAAGTTTCTGATGTTGACATGGGGCTCGCGAAACAGTATCGGCCACACATGATTAT	792	
Qy	2796	tttacagatccagatgtaccagctctcaggttgccttcattgattaccacagagacctcta	2855	
Db	793	TTTACAGATCCAGATGTACACAGCTCTCGAGGTCTCTCATGATTACACAGACCTCTA	852	
Qy	2856	caatcacctggagagcagcagatcgcccccagcctctacgtgccccctgggtccctcac	2915	
Db	853	CAATCACCTGGAGGACGACGAGATCGGCCCGCCAGCCTACGCTGCCCCCTGGTTCCTCAC	912	
Qy	2916	catgttgcctcacagttcccgctgggattgctagccagagcttctgattgattttct	2975	
Db	913	CATGTTTGCCTCACAGTTCGCCGTGGGATTCGTAGCCAGAGCTCTTTCATATGATTTTCT	972	
Qy	2976	tcagggaacagagctatatttaagtgctttaagctcttgggaagcctaaagccctt	3035	
Db	973	TCAGGGAAACAGAGTCAATTTAAAGTGGCTTTAAGTCTGTGGGAAGCCATTAAGCCCTT	1032	
Qy	3036	gattctgcagatgaacaccttagaaccatagttgactttataaaagcagcgtcaccaa	3095	
Db	1033	GATTCTGCAGCATCAAAACCTTAGAAACCATAGTTGACTTTTATAAAAGCAGCTACCCAA	1092	
Qy	3096	ccttggttggtacagatggaagagaccatacatcaggtatttgaaatgacatcgctaa	3155	
Db	1093	CCTTGGCTTGTACAGATGGAAAGACCATCAATCAGGTATTTGAAATGGACATCGCTAA	1152	
Qy	3156	acagttacaagcttatgaagttgagtcaccagctccttcaagaagaacttatcgattcttc	3215	
Db	1153	ACAGTTACAAGCTTATGAAGTTGAGTACCAGTCTCTCAAGAGAACTTATCGATTCTCTC	1212	
Qy	3216	tcctctcagtgacacccaagaatggataattagagaaacccaacagcagcttacgcaa	3275	
Db	1213	TCCCTCAGTGACACCAACCAAGATGGATAAATAGAACAGCAACACAGCAGCTTACGCAA	1272	
Qy	3276	acagaaccttgacctccttgaacagttgagcaggtgagcaggtggaagcttga	3335	
Db	1273	ACAGAACTTGAACCTTGAACAGTTCAGAGTGGCAAAATGGTAGAGTCAAAAGCCCTTGA	1332	
Qy	3336	ggccaccattgagagctcctcagcagtgagacgaacgaagcagcagcctgcttacct	3395	
Db	1333	GGCCACCATTTGAAAGCTCTCTGACATGAGAGCAAGCTGAGCAGGCCATGCTTACTCTT	1392	
Qy	3396	agaactggagcgttcgcccctgctgcagacggttgagagcgtcgcgugcgagcgcaga	3455	
Db	1393	AGAACTGGAGCGGTGCGCCCTGCTGCAGACGGTGGAGAGCTGGCGGCGGAGCGCAGA	1452	
Qy	3456	gccagcagccggagcgtgactgacagcagccgagcccgagcggcgactgacagcgtcg	3515	
Db	1453	GCCAGCAGCCGGAGCGCTGAGTGCAGCGACGCCGCGAGCCCGGAGCTGACAGCTCG	1512	
Qy	3516	caggagagattgcaacaccatccacacgtgccagcctt	3555	

Db 1513 CAGGAGAGATTGCAACACCATCCACACTGTCTCCAGGCCTT 1552

RESULT 9

AB011175

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

1624 a 1408 c 1432 g 1458 t

AB011175

Homo sapiens mRNA for KIAA0603 protein, complete cds.

AB011175

1

GI:3043729

Homo sapiens male brain cDNA to mRNA, clone_lib:pBluescriptII SK plus clone:HG1488b.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 5922)

Ohara, O., Nagase, T. and Ishikawa, K.

Direct Submision

Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute, DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnaif@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)

2 (sites)

Nagase, T., Ishikawa, K., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.

Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro

DNA Res. 5 (1), 31-39 (1998)

98290545

Location/Qualifiers

1. 5922

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HG1488b"

/sex="male"

/tissue_type="brain"

/clone_lib="pBluescriptII SK plus"

348. 4247

/gene="KIAA0603"

/gene="KIAA0603"

/codon_start=1

/product="KIAA0603 protein"

/protein_id="BAA25529.1"

/db_xref="GI:3043730"

/translation="MEPPSCIDPEPFPLEPGVSAQPGKPKSKRFLRWYVGGSCLDHRTTLPMLMAEIRRSOKPEAGGCGAPAAREVILVLSAPFLRCVPAPGAGSGTSPATOPNPAVFIFEKAOHSIFEHNSHDLTYFAVLIKAOPDDPSOMACHYFRADTSPQVDPVSIISIRQLSKAMKEDAKPSKDNEAFYNSOKFEVLKCGVTTHKAPSLIDDCMEKFFSLHEQRLKIQEGRGPDGEDADLEVVYVPGSPDCLPEADGDTYHLGLPAGSOPALTSRVCFFPERILSDSGFDEQEQFRCSSVTVQRRVHBSQSKQPRRHASAPSHVQPSDSKRNMTLQVGFREINLISPTKSVVLEKFNKDISCSQGIKHVDFHGFICRESPEGLSQICYVFCQASSELVDEMLTKQAFSTAAALQSAKQIKLCEACPMHSLKLCERIEGLYPRAKLTQRLSLTQNEQADIFERVQKMPVSDQENELVILHRLQCEAKQKTHVHIGEPSTISNSTIPENATSSGREFKLDILKNKRRSLSSLENIISRGANRMRLGSDVSDFERNLSASEKDYSPGDSPPCTPPASPPSSAQVTPPEEDSPQFRRRAHTFSHPSTSRKLNLDQRAQGVRSPLLRQSSQESQSNLSSVRMYKESNSSSLSLTFSFAPSFTSFLKSFYQNSGRSLSPQYENIETQDTSASDQEGKRTSSTCSNESLVSQSVTPPRISWRQIRFLRVASPMNKSPSAMQOQDGLDNELLPLSPTMEEPLVIFLSDGDPKIEERKSLRWLKAHQITLLLRMEKXENOKLEGASRDELOSKVKLDYEVGACQEVLTMDKLLNCRKLRCDMEDLHTLKREGVPSRGEIWOFLALQYLRHLRNKQPPDISYKELLKQLTAQHALLVDLGRFTPTHPYSVOLGQGLSLNLLKAYSLLDKEVGYCGQISFVAGVLLHMSSEQAFEMKFLMYDLGFRKRYQDMSLQIQMYQLSRLLHDYRDLNHLNEEISPSLYAAPFLTLFASQFSLGFAVRFDIIFLQGTIVFKVALSLSSQETLIMECSFENVEFLKNTLPDMNTSEMEKIIITQFEMDISKOLHAYEVHYVLQDELQESSYSCSESETELEKRNQSLKRNQMDLLEKLYAHKTIQALSNLENLLTRETMMKSLIRTLEQEKWAYOKTVEOLRKLFPADALANCDLLRLDNCNPNKAKIGNKP"

	Query Match	19.2%;	Score 582.2;	DB 9;	Length 5922;
	Best Local Similarity	56.6%;	Pred. No. 1.9e-138;		
	Matches 1585; Conservative	0;	Mismatches 1028;	Indels 186;	Gaps 10;
Qy	823 agcggacacaatattgtcacgccacagatatcgaggaaatacgaaactatctcttcacg	882			
Db	1368 AGCGACCCAGTCAGCTCCAGCCTCGGACTCGGAGATCGGAGAACAAGACAGGACCATTGCTCTTCCAG	1427			
Qy	883 attggccagtctgaagtttacctcatcagtccttgacacacaaaaaatagcattggagaaa	942			
Db	1428 GTTGGGCATTGTGAGATTAACTTTATCATAGTCCAGACATAAATAGTTGTGCTAGAAAAAG	1487			
Qy	943 aattttaaggagatatccctttgtctctcagggccatcagacgtagaaccaatttggggtt	1002			
Db	1488 AATTTTAAGATATCTCTCTGTGTCTCAGGGTATAAAGCATGTGATCACTTTGGCTTT	1547			
Qy	1003 atctgtcggggagcttcgcggaggtggcgcttccattttgtctgttaagtgttttcagtcg	1062			
Db	1548 ATCTGCCGGAGTCTCCAGAGCCTGGACTTAGCCAGTATAATTGTTATGTATTCCAGTGT	1607			
Qy	1063 acaaatgaggtctcgttgtgatgaattatgatgacctgaaacaggccottcacggtggcc	1122			
Db	1608 GCCAGCGAATCTCTGGTGTGATGTAGGAATATGCTGACTCTGTAACAGGCCCTTCAGTACGGCG	1667			
Qy	1123 gcagtgcagcagacagctaaagcgcagcccagctgtgtgagggctgcgccctgcacaaagc	1182			
Db	1668 GCTGCCCTGCAGATGCCAACAGCAGATTAACTGTGTGAGGCCCTGCCCGATGCACCTT	1727			
Qy	1183 ctgcacagctctgtgagaggatagagggaatgaattcttccaaacaaaaactagaactg	1242			
Db	1728 TTGCATAAAGCTCTGTGAAGAGATTGAAGGTCTTACCCACCAAGAGCCAAGCTGGTGATA	1787			
Qy	1243 caaagcacctgcagcaattaaccaatcagcagcagcgactatttttgaagaggttcag	1302			
Db	1788 CAGAGGCATCTCTCATCTACTCACAGATAATCAGCAAGCTGCACATCTTTGAAGAGATTCCAG	1847			
Qy	1303 aaattgagccgagaatgagcagcgagagaatgaattgattattcttctctcagatgt	1362			
Db	1848 AAAATGAAGCCAGTCAGTGACCAGGAAGAAAATGAACCTTGTGATTTTACACTGAGGCAG	1907			
Qy	1363 ttaatgaagagaacaagaaagacacatccatttggggagatgaagcagacatcgacg	1422			
Db	1908 CTGTGTGAAGCCAGCAGCAAGACACATGTCACATCGGG-----AAGGCCCTTCT	1958			
Qy	1423 atggcagcagagaattattggaagtgaattaccaccocagctgcactcgatttagcctagat	1482			
Db	1959 ACTATTTCAAATAGTACAATCCGAGAAAATGCAACAGCAGTGGGAGTTTCAAACTTGAC	2018			
Qy	1483 atgtgaaaaaacaagaaagagatctttaaagagtcctttagaagaatattttgtccogg	1542			
Db	2019 ATTCTGAAAAATAAGCTTAAGAGATCTCTTAAGTCTCCCTGGAAAAATATCTTCTCAAG	2078			
Qy	1543 gg--taataaagccagagccttcagggaacactccactgag-----1581				
Db	2079 GGAGCTAACAAATGAGAGGTCGGCTTGGAGTGTGGACAGTTTTTGAACGGTCCACACAGT	2138			
Qy	1582 -----1581				
Db	2139 CTTGCTTCAGAGAAGGACTACTACACAGGGATTCTCCACCAAGGAGACACCGCCAGCGTCC	2198			
Qy	1582 -----gtgactctgtagctcctctcagtt	1608			
Db	2199 CCACGGTCTCAGCTTGGCAAAAGTTTCCCAGAGAGGATTCCGACTCCCCGAGTTTCGA	2258			
Qy	1609 acattaaagtaacaccagcaagagagccattctgtgtgaaaaggagccttggccactctct	1668			
Db	2259 AGACGGGCACACAGCTTCAGGCCACCCACCTTCAAGCACAAAGAGAAAGCTGAATTTGCGAC	2318			
Qy	1669 gagagctcctttaagctcctcgtcctcgtg-----gagg	1702			
Db	2319 GATGGGAGGGCTCAGGGGTGCGTGTCCCTCTGCTGAGCGCAGAGCTCCAGTGAACAGTGC	2378			

Qy	1703	acctgtccagtactcggagagtcattctcccagaagaagaccagctccgctgtgtgcgcccaagc	1763
Db	2379	AGCAATCTTTTCGTACGTTGCACGCATGTACAAGAGAGTAATTTCTTCTCCAGCTTTCCA	2438
Qy	1763	aggcttcicaggagcgagcaaacacacccttagtcaacttccccatcgaa'tgcagagaacctc	1822
Db	2439	AGTCTTCACACTTCTTCTGTGCCCTCCTTCACTGCGCCCTCTTTCTTGAAAAGCTTT	2498
Qy	1823	cacaacctgcccgggggtc---cccgggggttttcgcaaaggaaaacttatgagggatacact	1879
Db	2499	TACCAGAAATTTCAGGTAGACTGTCCCACAGTAGTATGAATAATGAATCAGACAAGACACTGCT	2558
Qy	1880	cagt'gagcacagagagcctcatgaacgaagaagactt'gaatccaagaacaaacctcttg	1939
Db	2559	TCAGAA'TCAAGTGATGAGGAAGGGAGAAAGAAGCCCTCATCTTACCTGCAGCAANTGAGTCC	2618
Qy	1940	gtgatct'-ggtggactcctgtgaagaccccggaggagcattctctggagcagcagatatctc	1998
Db	2619	CTAAGTGTGGGAGGAACCTCTGTCACTCTCGCGGATCTCTCTGGCGCAGCGCATTTTC	2678
Qy	1999	ctcgcagtagcccccocgacagaaggcgtgcgattcttccagcagatat'gaagattattca	2058
Db	2679	CTCAGGGTGTCTTCTCCAT'GAAACAATCTCCCTCAGCAA'TCAACACAGAGA'TGGAT'GTG	2738
Qy	2059	gagctgggagagcttcccacagatctctttagaacactgttgtgaagatggggccctt-	2117
Db	2739	GACAGGAACGAGCTGTGCGCACCTGCTCCCCCTCTCTCCAACCATGGAGGAGAACCGCTG	2798
Qy	2118	-----tgcccccccccacagaggagaaaaagaaagaaagacatctcgt	2154
Db	2799	GTTATATTCCTGTCTGGGAGGATGACCCAGAAAAAGATT'TGAAGAAGAAAGAAATCAAAA	2858
Qy	2155	gagctcgcagagctgtgcgaaggctattottcaacagatactctcttagaatggag	2214
Db	2859	GAATCTGAGGAGCTGTGTGGAAAGACTATACACAAACAATCTTGTACTTCGAATGGAA	2918
Qy	2215	aaggaaaaatcagaagctcca- --agcctctgaaaaatgttgtgtaacaagcgcctgaag	2271
Db	2919	AAAGAAAAACCAGAAACTTGAAGGAGCAACGAGATGAACTCCAGTCCACAGAAAAAGTTAAA	2978
Qy	2272	ctcgattatgaagaaatactccctgtctttaagaagaatactaagttgggaaagatg	2331
Db	2979	TTAGACTATGAAGAAGTTGGTGCATGTCAGA'AAGAGGTCTTAA'TAAC'TTGGATTAAGAAG	3038
Qy	2332	cttagcactccaggaagatcaaaaa'taagtttgacatggaaaaaatcagctcgctgtt	2391
Db	3039	TTGTITAA--ACTGCAGAGCTAAATCAGATGTGATATGGAAGATATTCATACTCTTCTT	3095
Qy	2392	gggcaaggtgtccaoigtcatcaccaggtgaaatctcggaaattctagctgagcaattc	2451
Db	3096	AAAGAAGAGGTCTCCAAAAGTCCACAGAGAGAAAT'TTGGCAGTTTCTGGCTTTACAGTAC	3155
Qy	2452	cacctaaacacacagtttcccagcaaacacgagcgcacaaaggatgtgccatacaagaacctc	2511
Db	3156	CGACTCAGACACAGATTGCCCTATNAACAAACAGCCCTCGTACATATCTATAAAGAACTT	3215
Qy	2512	ttaaagcagctgacttcccagcagcatcgatcttattgaccttggcggaacctttcct	2571
Db	3216	TTGAAGCAGCTCACTGTCTACGACAGCATGGATTCGTGTGGATTTAGGAAGACCGTTTCTCT	3275
Qy	2572	acacaccatactctctgccagcttgagcagacagcattcgctttacaacatttg	2631
Db	3276	ACTCACCTTTACTTTTCAGT'TACAGCTTTGGCCAGGNACAGCTGCTCACTGTTTAACTCTCTG	3335
Qy	2632	aaggcctactcaactcttagaccaggaagtgggatatattgccaaggtctcagcttttgtagca	2691
Db	3336	AAAGCCTATTCTTTGTGGACAAAAGATGGGATACTGT'CAGGGGATCAGCTTTTGTGGCT	3395
Qy	2692	ggcattttgtcttctcata'tgadtgagggaagggttttaaaatgctcaaatgttctag	2751
Db	3396	GGAGTCTGCTTCTGCACT'GATGATGAAGAGCAAGCCTTTTGAATGCTGAATTTCTCTATG	3455
Qy	2752	tttagcatgagggctgcgcaaacagtatcagccagacatgattattttacagatccagatg	2811


```

Db 3456 TATGACCTCGGCTTCGCGAAGCAGTACAGACCTGACATGATGTCGCTCGATTCACAAATG 3515
Qy 2812 Taccagctctcaggttcttctcattaccacagagacacctctacaatcacctggagag 2871
Db 3516 TACCAGCTGTCCAGGCTCTCTCATGACTATACAGAGATCTCTACATCACCTTGAAGAA 3575
Qy 2872 caccagatcgcccccagcctcagctcagctcgcctccctcctcctcctcctcctcag 2931
Db 3576 AATGAAATCAGCCCGCAGTCTTATGCTGCCCTGGTTCCTCACATTTGTCCTCTCAG 3635
Qy 2932 ttcgcgtggatctcagcagagctcttcttatgatatttttttcttcagggaacagagtc 2991
Db 3636 TTTTCATTAGGATTTGTAGCAGAGTCTTTTGTATATTTTCTTCAGGAACTGAAGTT 3695
Qy 2992 atatttaagtggctttaaagtctgtgggaagcagataagccttgattctcagcatgaa 3051
Db 3696 ATATTCAAGTTGCACTCAGCCTACTGAGCAGCCCAAGAGACACTATATGAATGTGAG 3755
Qy 3052 aacctagaaccatagttgacctttataaaagcagcgtacctcaaaccttggcttgtaacag 3111
Db 3756 AGCTTTGAAATATTTGTGAGTTCTTAAACACAGCTTACCTGATGATGAATACCTCTGAA 3815
Qy 3112 atggaagaccatcaatcaggtatttgaatgagacatcgctaaacagttacaagcttat 3171
Db 3816 ATGGAATAAATATTACCCAGGTTTTTGAGATGGATATTTCTAAGCAGTTGCATGCCCTAT 3875
Qy 3172 gaagttgagtcaccagctcttcaagaagaacttatcgattctctctcctcctcagt---gac 3228
Db 3876 GAGGTGGAATATCATGTCTACAGGATGAGCTTCAGGAATCTTCATATTCCTGTGAGGAT 3935
Qy 3229 aaccataaagatggataattagaaaaaaccaacagcagcttacgcacaaacagacacttgac 3288
Db 3936 AGTGAACATTTGGAGAACCTGGAGAGGCCCAATAGCCCACTGAAAGACAAACATGAC 3995
Qy 3289 ctccctgacagttcaggtgagcaaatggttaggtatcccaagccttgaggccaccattgag 3348
Db 3996 CTCCTAGAAAATTTACAGGTAGCTCTACTAATAATCCAGGCTTGGAAATCAAACTGGA 4055
Qy 3349 aagctcctgagcagtgagcaagctgaagcagcagcagccttaccttgaactggagcgg 3408
Db 4056 AATCTTTTGACGAGAGACCAAAATGAAGTCTTTAATCCGACCCCTGGAACAGAAAAA 4115
Qy 3409 tcggccctcgtcagacagctgagagagctgagcggcg 3447
Db 4116 ATGGCTTATCAAAAGACAGTGGAGCACTCCCGAAGCTG 4154

RESULT 10
AC009595/c standard; DNA; HTG; 189811 BP.
XX AC009595;
XX AC009595;
SV AC009595.3
XX 30-AUG-1999 (Rel. 60, Created)
DT 03-JUL-2000 (Rel. 64, Last updated, Version 4)
DE Homo sapiens chromosome 4 clone RP11-392K14 map 4, WORKING DRAFT SEQUENCE,
DE 28 unordered pieces.
XX HTG; HTGS_DRAFT; HTGS_PHASE1.
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX [1]
RN 1-189811
RP Birren B., Linton L., Nusbaum C., Lander E.;
RT "Homo sapiens chromosome 4, clone RP11-392K14";
RL Unpublished.

```

```

XX RN 1-189811
RP Birren B., Linton L., Nusbaum C., Lander E., Allen N., Anderson M.,
RA Baker J., Baldwin J., Barna N., Beckerly R., Benn J., Brown A., Castle A.,
RA Cerny J., Colangelo M., Collins S., Collymore A., Cooke P., Dearellano K.,
RA Depayre E., Devon K., Dewar K., Donelan L., Doyle M., Ferreira P.,
RA Fitzhugh W., Forrest C., Funke R., Gage D., Galagan J., Gardyna S.,
RA Gilbert D., Grant G., Hagos B., Heaford A., Horton L., Howland J.C.,
RA Jones C., Kann L., Karatas A., Lehoczy J., Lieu C., Locke K.,
RA McDonald P., Marquis N., McEwan P., McGurk A., McKernan K., McLaughlin J.,
RA Melford J., Molla M., Morris W., Morrow J., Mychaleckyj J., Naylor J.,
RA Niloff M., O'Connor T., O'Donnell P., Pavlin B., Peterson K., Pollara V.,
RA Riley R., Roberts D., Roy A., Severy P., Stange-Thomann N., Stojanovic N.,
RA Stone C., Subramanian A., Tesfaye S., Torruella-Miller I., Vassiliev H.,
RA Vo A., Wagner A., Wheeler J., Wu X., Wyman D., Ye W.J., Zody M.;
;
Submitted (27-AUG-1999) to the EMBL/GenBank/DBJ databases.
Whitehead Institute/MIT Center for Genome Research, 320 Charles Street,
Cambridge, MA 02141, USA
On Jun 23, 2000 this sequence version replaced gi:6006228.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1755
Center clone name: 392_K14
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-primer-amersham; 5% of reads
Chemistry: Dye-terminator Big Dye; 95% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 164704 bases at least Q40
Consensus quality: 175594 bases at least Q30
Consensus quality: 181024 bases at least Q20
Insert size: 214000; agarose-fp
Insert size: 187111; sum-of-contigs
Quality coverage: 3.0 in Q20 bases; agarose-fp
Quality coverage: 3.5 in Q20 bas.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 451: contig of 451 bp in length
* 452 551: gap of 100 bp
* 552 2663: contig of 2112 bp in length
* 2664 2763: gap of 100 bp
* 2764 4124: contig of 1361 bp in length
* 4125 4224: gap of 100 bp
* 4225 6767: contig of 2543 bp in length
* 6768 6867: gap of 100 bp
* 6868 8750: contig of 1883 bp in length
* 8751 8850: gap of 100 bp
* 8851 11272: contig of 2422 bp in length
* 11273 11372: gap of 100 bp
* 11373 13411: contig of 2039 bp in length
* 13412 13511: gap of 100 bp
* 13512 16962: contig of 3451 bp in length
* 16963 17062: gap of 100 bp
* 17063 22544: contig of 5482 bp in length
* 22545 22644: gap of 100 bp
* 22645 26284: contig of 3640 bp in length
* 26285 26384: gap of 100 bp

```

```
CC * 26385 31223: contig of 4839 bp in length
CC * 31224 31223: gap of 100 bp in length
CC * 31224 36014: contig of 4691 bp in length
CC * 36015 36114: gap of 100 bp
CC * 36115 41939: contig of 5825 bp in length
CC * 41940 42039: gap of 100 bp
CC * 42040 48122: contig of 6083 bp in length
CC * 48123 48222: gap of 100 bp
CC * 48223 52983: contig of 4761 bp in length.
CC * 52984 53083: gap of 100 bp
CC * 53084 58094: contig of 5011 bp in length
CC * 58095 58194: gap of 100 bp
CC * 58195 64606: contig of 6412 bp in length
CC * 64607 64706: gap of 100 bp
CC * 64707 71079: contig of 6373 bp in length
CC * 71080 71179: gap of 100 bp
CC * 71180 79886: contig of 8707 bp in length
CC * 79887 79986: gap of 100 bp
CC * 79987 86155: contig of 6169 bp in length
CC * 86156 86255: gap of 100 bp
CC * 86256 95370: contig of 9115 bp in length
CC * 95371 95470: gap of 100 bp
CC * 95471 103712: contig of 8242 bp in length
CC * 103713 103812: gap of 100 bp
CC * 103813 113565: contig of 9753 bp in length
CC * 113566 113665: gap of 100 bp
CC * 113666 123867: contig of 10202 bp in length
CC * 123868 123967: gap of 100 bp
CC * 123968 132296: contig of 8329 bp in length
CC * 132297 132396: gap of 100 bp
CC * 132397 141110: contig of 8714 bp in length
CC * 141111 141210: gap of 100 bp
CC * 141211 161622: contig of 20412 bp in length
CC * 161623 161722: gap of 100 bp
CC * 161723 189811: contig of 28089 bp in length.
XX
FH Key Location/Qualifiers
FH source 1. .189811
FH /chromosome="4"
FH /db_xref="taxon:9606"
FH /organism="Homo sapiens"
FH /map="4"
FH /clone="RP11-392K14"
FH /clone_lib="RP11-11 Human Male BAC"
FH misc_feature 1. .451
FH /note="assembly_fragment clone_end:T7 vector_side:right"
FH misc_feature 532. .2663
FH /note="assembly_fragment"
FH misc_feature 2764. .4124
FH /note="assembly_fragment"
FH misc_feature 4225. .6767
FH /note="assembly_fragment"
FH misc_feature 6868. .8750
FH /note="assembly_fragment"
FH misc_feature 8851. .11272
FH /note="assembly_fragment"
FH misc_feature 11373. .13411
FH /note="assembly_fragment"
FH misc_feature 13512. .16962
FH /note="assembly_fragment"
FH misc_feature 17063. .22544
FH /note="assembly_fragment"
FH misc_feature 22645. .26284
FH /note="assembly_fragment"
FH misc_feature 26385. .31223
FH /note="assembly_fragment"
FH misc_feature 31324. .36014
FH /note="assembly_fragment"
FH misc_feature 36115. .41939
FH /note="assembly_fragment"
FH misc_feature 42040. .48122
FH /note="assembly_fragment"
```

```
FT misc_feature 48223. .52983
FT /note="assembly_fragment clone_end:SP6 vector_side:left"
FT misc_feature 53084. .58094
FT /note="assembly_fragment"
FT misc_feature 58195. .64606
FT /note="assembly_fragment"
FT misc_feature 64707. .71079
FT /note="assembly_fragment"
FT misc_feature 71180. .79886
FT /note="assembly_fragment"
FT misc_feature 79987. .86155
FT /note="assembly_fragment"
FT misc_feature 86256. .95370
FT /note="assembly_fragment"
FT misc_feature 95471. .103712
FT /note="assembly_fragment"
FT misc_feature 103813. .113565
FT /note="assembly_fragment"
FT misc_feature 113666. .123867
FT /note="assembly_fragment"
FT misc_feature 123968. .132296
FT /note="assembly_fragment"
FT misc_feature 132397. .141110
FT /note="assembly_fragment"
FT misc_feature 141211. .161622
FT /note="assembly_fragment"
FT misc_feature 161723. .189811
FT /note="assembly_fragment"
XX
SQ Sequence 189811 BP; 52455 A; 40839 C; 40321 G; 53485 T; 2711 other;

Query Match 13.1%; Score 465.8; DB 30; Length 189811;
Best Local Similarity 99.6%; Pred. No. 5.2e-91;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 416 aadgcctgagatcatcctccatccgtcagcgccggggaagatcgccgcagagagc 475
DB 97313 AGGTGCTGAGATCATCAGCTCCATCCGTGAGCGGGGAAGATCGCCGCGAGGAGC 97254
QY 476 tgcactgccgcgtccgaggttcgacgacacgttttccaaagattcgagggtctttctgcg 535
DB 97253 TGCAC'TGCCCGTCCGAGTTCGACGACACGCTTTTCCAAAGATTTCGAGGTCTCTTCTCGG 97194
QY 536 gccgcgtgacgtggtgcgcacaagaagctccgcgcctcgatcgacgagtcgacgaga 595
DB 97193 GCGCGGTGACGTTGGCGCACAGAGGCTCCGCGGCCCTGATCGACGAGTGTGATCGAGA 97134
QY 596 agttcaatcacgtcagcgagccgagcgccgggttcgagagagcccccccaaccccccccatg 655
DB 97133 AGTTCAATCACGTGACGCGGCGCGGGGTCCGAGAGCGCCCGCCCAACCCGCCCATG 97074
QY 656 ccgcgccccacagggaccagagagcctgtgcgagggccccatgcgaagtctcttccagc 715
DB 97073 CCGCGCCCAACAGGGAGCCAGGAGCCTGTGCGCAGGGCCCATGCGCAAGTCTCTTCCAGC 97014
QY 716 ccggcctgcgctcgctggccttttaggaagagctgcagagatgggggctccgaagcagcg 775
DB 97013 CCGGCGCTGCGCTCGCTGGCCTTTAGGAGAGAGCTGAGGATGGGGCTCCGAAGCAGCG 96954
QY 776 gcttctcagctccttcgagggagcgcacattgagaacacacctattagcgggacacaata 835
DB 96953 GCTTCTTCAGTCTCTTCGAGGAGAGCGACATTGAGAACCACCACTCATTTAGCGGACACAATA 96894
QY 836 ttgtgcagccacagatatcgaggaaaaatcgaaactatgctcttcacgat 884
DB 96893 TTGTGCAGCCCAACAGATATCGAGGAAAAATCGAACTATGCTCTTTCACGCT 96845

RESULT 11
AC021106 AC021106 195108 bp DNA linear PRI 03-JUL-2001
LOCUS Homo sapiens clone RP11-177C12, complete sequence.
DEFINITION
```


[illegible]

Qy	2211	ggagagggaaaaatcagaagctccaagcctctgaaaaatgatttgctgaaacaagcgccctgaa	2270
Db	2487	ggagaaggagaatcagaagctacaagcctctgaaaacgatTTGCTGAACAACGCCCTCAA	2546
Qy	2271	gctcgattatgaagaaattactccctgctctaaagaagtaactacagctggtggaaagaat	2330
Db	2547	GCTTGACTATGAAGAAATCACTCCGTGCTTTAAGAAGTCACTACAGTGTGGGNAAGAT	2606
Qy	2331	gcttagcactccagggaagatcaaaaattaagtcttgacatggaaaaaatgcactcgctgt	2390
Db	2607	GCTTAGCCTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAAGTGCACCTCAGCTGT	2666
Qy	2391	tgggcaagctgcccagctcatcacccaggtgaaattctgaaattctgaaattctactcaagaatt	2450
Db	2667	TGGCAAGGTGTGCCACGTCACTACCGAGGTGAGATCTGGAAATTTCTAGCTGAGCAGTT	2726
Qy	2451	ccacctaaacaccagittccccagcaaacagcagccaaaggatgctcatacaaaagaact	2510
Db	2727	CCACCTTAACACCCATTTCCCTAGTAAACAGCAGCAAGGACGTGCCCTACAAAGAGCT	2786
Qy	2511	cttaagaagctgaactcccagcagcatgcgattcttatgaaccttggggagaacctttcc	2570
Db	2787	CTTGAAGAAGCTGACCTCGCACAGCAGCGCCATTCTCATCGACCTCGGGCGAACCCTTTC	2846
Qy	2571	tacacaccatactctctgcgccagcttggagcagcagctatcgtcttacaacattt	2630
Db	2847	AACACATCCATACTCTCTGCCACCTGGAGCAGGTCAGCTGTCACTTTACAACATTC	2906
Qy	2631	gaaggcctactcactctagacccaggaagtgggatatgtccaaaggtctcagcttgtagc	2690
Db	2907	GAAGGCCTACTCGCTCTCGACAGAGAGGTTGGATCTATGCCAAGGCTCAAGCTTTGTGGC	2966
Qy	2691	aggatttgcctctctcatatgagtgaggaagagcgctttaaagtctcaagttctcgt	2750
Db	2967	AGCATTTTGCTTCTTCACATGAGTGAGGAAGAGCGGTTCAAGATGCTCAAGTTCCTGAT	3026
Qy	2751	gtttgacatggggctgcggaaacagtatcggccagacatgattatttcacagatccaagt	2810
Db	3027	GTTTGATATGGGGCTCGGAAACAGATATCGGCCACAGCATGATTTATTGTCAGATCCAAG	3086
Qy	2811	gtaccagctctgaggtgcttcatgattacacagagacctcacaaacctcctcgagga	2870
Db	3087	GTACAGCTGTACGGCTCCTCCACGATTCACCGAGACCTCTCAACACACCTGGGAAGA	3146
Qy	2871	gcacgagatcgcccacagcctcagctgcgcccttggtctcaccatgtttgctctaca	2930
Db	3147	GCACGAGACTGGCCCCCTACCTACGCGGCTCCCTGGTTTCTCACCGGTGTGCGCTCACA	3206
Qy	2931	gttccgcgtgggattctgacagagtttggatatgattttcttcagggaacagaggt	2990
Db	3207	GTTCCACATCGGGCTTGTAGCCAGAGTCTTTGATATGATCTTCCTCAGGGATCAGAGGT	3266
Qy	2991	catatttaaagtggctttaaagctgtgtgggaagccataagcccttgattcttcgagcatga	3050
Db	3267	CATATTTAAGTAGCTTTAAGTCTTTGGGAGGCCATAAGCCCTTGATTCTACAGCATGA	3326
Qy	3051	aaacctagaaccatagttgactcttataaaaagaacgctacccaaaccttggcttggtaca	3110
Db	3327	GAACCTTGGAAACCATCGTGAGCTTCATAAAGAACACACTCCCAACCTGGGCGCTGTGCA	3386
Qy	3111	gatgaaagagaccatacaacaggtattgaaatggacatcgctaaacagttacaaagctta	3170
Db	3387	GATGGAGAAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCAAGCAGCTCCAGGCCTA	3446
Qy	3171	tgaagttgagtaccacgtccttcaagaagaacttatcgattctctctcctcctcagtgacaa	3230
Db	3447	TGAGGTCGATACACGTCTGTCCAGGAGGAGCTTATTGAGTCCCTCGCCTCTCACTGACAA	3506
Qy	3231	ccaaagaatgataaattagacaaacacagcagcttacgcaaaacagacacttgacct	3290
Db	3507	CCAAAGAAATGGAAATTTGGAGAAACCAACAGCAGCTTGGCGCAACAGAACCTTTGACCT	3566
Qy	3291	ccttgacagcttgacgttgccaaatggttaggtatccaaagcctctgaggccacacttgagaa	3350

```

Db 3367 CCTGGAGCAGTTGCAGGTGCAGAAATGCTTAGGATCAAAAGCCTTGAAGCCACCGTAGAGAA 3626
Qy 3351 gctctgagcagtgagagcaagctgaagcaggccatgcttaccttagaactggagcggtc 3410
Db 3627 ACTTCATTACCAGCGAGAGTAAGCTGAAGCAGCGTGCCTGACCTGGAGGTGGAGCG--T 3684
Qy 3411 ggcctctgctcagacggctgagagagctgcggcggagagcgagagccagccagccagcgga 3470
Db 3685 CGCCTCTCTCAGATGTTGGAGGAGTGTGGCGAGGCGGAGGCAAAAGCGCCGCGCCAGCACTCCAGA 3744
Qy 3471 gcttgagtcacacgacccgagcccgagcccgagcgagcgactgacagct 3512
Db 3745 GCCAGACTGCCACCAGCTGGAGGCCACAGCGGATTCACCGCT 3786

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ptzgpt-F1s
US-08-232-463-14

```

Query Match 1.68: Score 57.8; DB 1; Length 7218;

LENGTH: 9515 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical isolate P2-2
US-08-920-827-13

Query Match 1.38; Score 45.4; DB 1; Length 9515;
Best Local Similarity 46.28; Pred. No. 0.033; Indels 0; Gaps 0;
Matches 151; Conservative 0; Mismatches 176;
Qy 438 catcgtcaggcgggaagatcgcggcaggagagtgctcactgcgcgtccgagttcga 497
Db 802 CATGGCAAGCGGTGATGGAGCCTGGAAATCATGATACCGCGCGCCACGCTCT 861
Qy 498 cgacacgtttccaaagaagttcgggtgtcttcttgcggcggcgtgacggtggcgacaa 557
Db 862 CGCCTGGTATCGGAAGCCCTGACAAAGCTCTACGACAGGTGCGCGCGCGCCAGCA 921
Qy 558 gaaggtcgcggccctgatcgacagtgatcgagagagttcaatcacgtcagggcag 617
Db 922 GACCCGTGCCACCATTAACCCGCGTGCCTGGGGGTGATCGCGCGGTGGTGGCGTGA 981
Qy 618 ccgggggttcgagagagcccccgcacaccccgcccgcccgcccgcccgcccgccgga 677
Db 982 CTTCCCGCTCGACATGCGCGCTGGAGCTCGCCCGCGCCCTGGCGCGCAACTCGT 1041
Qy 678 gcctgtgcgagggcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgccgtt 737
Db 1042 GGTGCTCAAGCGCGCGCGAGTATCGCGCTTCTCGCGCTGCGCGTGGCGAGCTGGCCCT 1101
Qy 738 taggaagagctgcaggtgagggcct 764
Db 1102 GGAGCGGGGTGCGGAGGCGGTGCT 1128

RESULT 5
US-08-921-177-13
Sequence 13, Application US/08921177
Patent No. 5798211
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921.177
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 9515 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical isolate P2-2
US-08-921-177-13

Query Match 1.38; Score 45.4; DB 1; Length 9515;
Best Local Similarity 46.28; Pred. No. 0.033; Indels 0; Gaps 0;
Matches 151; Conservative 0; Mismatches 176;
Qy 438 catcgtcaggcgggaagatcgcggcaggagagtgctcactgcgcgtccgagttcga 497
Db 802 CATGGCAAGCGGTGATGGAGCCTGGAAATCATGATACCGCGCGCCACGCTCT 861
Qy 498 cgacacgtttccaaagaagttcgggtgtcttcttgcggcggcgtgacggtggcgacaa 557
Db 862 CGCCTGGTATCGGAAGCCCTGACAAAGCTCTACGACAGGTGCGCGCGCGCCAGCA 921
Qy 558 gaaggtcgcggccctgatcgacagtgatcgagagagttcaatcacgtcagggcag 617
Db 922 GACCCGTGCCACCATTAACCCGCGTGCCTGGGGGTGATCGCGCGGTGGTGGCGTGA 981
Qy 618 ccgggggttcgagagagcccccgcacaccccgcccgcccgcccgcccgcccgccgga 677
Db 982 CTTCCCGCTCGACATGCGCGCTGGAGCTCGCCCGCGCCCTGGCGCGCAACTCGT 1041
Qy 678 gcctgtgcgagggcccgcccgcccgcccgcccgcccgcccgcccgcccgcccgccgtt 737
Db 1042 GGTGCTCAAGCGCGCGCGAGTATCGCGCTTCTCGCGCTGCGCGTGGCGAGCTGGCCCT 1101
Qy 738 taggaagagctgcaggtgagggcct 764
Db 1102 GGAGCGGGGTGCGGAGGCGGTGCT 1128

RESULT 6
US-08-362-577C-13
Sequence 13, Application US/08362577C
Patent No. 5807673
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362.577C
FILING DATE: 27-MAR-1995

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
;
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Pseudomonas aeruginosa
; STRAIN: Clinical Isolate P2-2
; US-08-920-828-13
;
; Query Match 1.3%; Score 45.4; DB 2; Length 9515;
; Best Local Similarity 46.2%; Pred. No. 0.033;
; Matches 151; Conservative 0; Mismatches 176; Indels 0; Gaps
;
; QY 438 catcctcaggcgggaagatcgcgcgcgcagagagctgcgaactgccgtccgagttcga 497
; Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 802 CATGGCAACCCGGTGATGACCCCTGGAAACATCATGTATCCCGCGCGCGCCACGCTT 861
;
; QY 498 cgacacgttttccaagaagtctgagtgctcttctgcgcgcgcgtgacgggtggcgcaaa 557
; Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 862 CGCTGTGATCGCAAGGCTCGACAACTCTACACACAGCTCTACGCGCGCGCGCCACGA 921
;
; QY 558 gaaggctcgcgcggccctgatcagcagtgcatcagaaagtccaatcaatcagcagcgag 617
; Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 922 GACCCTGGCCACCATTTACCCGCGCTGCGCGTGATCGCGCGCGGTGGTCCGTGGAA 981
;
; QY 618 ccggggggtccgagagccccgcgcacacccgcccatgccgcgcacagggagccagga 677
; Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 982 CTTCGCCGCTGCATGGCGCCCTGGAAAGCTCCGCCCGCCCTGGCGCGCACTCGGT 1041
;
; QY 678 gcctgtgcgagggcccatcgcaagtccttctccagccccggcctgcgtcgctggcctt 737
; Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 1042 GGTGCTCAAGCGCGCAGCAGTCGCGCTTCTCCGCCCTGGCGCTGGCGAGCTGGCCCT 1101
;
; QY 738 taggaaggagctcaggaatgggggacct 764
; Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
; Db 1102 GGAGGCGGGGTGCGGAAAGCGGTGCT 1128
;
; RESULT 8
; US-08-253-155A-9
; Sequence 9, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jenio
; APPLICANT: Draetta, Giulio
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA

```

```

, ZIP: 02109
,
, COMPUTER READABLE FORM:
, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: ASCII(text)
,
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/253,155A
, FILING DATE: 02-JUN-1994
, CLASSIFICATION: 435
, ATTORNEY/AGENT INFORMATION:
, NAME: Vincent, Matthew P.
, REGISTRATION NUMBER: 361,709
, REFERENCE/DOCKET NUMBER: MII-028
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (617) 227-7400
, TELEFAX: (617) 227-5941
, INFORMATION FOR SEQ ID NO: 9:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 8201 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: CDNA
,
, US-08-253-155A-9

```

Query Match	1.3%	Score 45; DB 1; Length 8201;
Best Local Similarity	53.0%	Pred. No. 0.039;
Matches 96: Conservative		0: Mismatches 85: Indels

```

1  RESULT          9
2  US-08-295-060-3
3  Sequence 3, Application US/08295060
4  Patent No. 5659123
5  GENERAL INFORMATION:
6  APPLICANT: VAN RIE, Jeroen
7  APPLICANT: JANSSENS, Stefan
8  APPLICANT: PERFEROEN, Matrix
9  TITLE OF INVENTION: NEW DIABROTIC
10 NUMBER OF SEQUENCES: 4
11 CORRESPONDENCE ADDRESS:
12 ADDRESSEE: Burns, Doane, Swedell &
13 STREET: P.O. Box 1404
14 CITY: Alexandria
15 STATE: Virginia
16 COUNTRY: United States
17 ZIP: 22131-1404
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/08/295
25 FILING DATE: 26-AUG-1994

```

```

CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 010830-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1957 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: YES
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1947
US-08-295-060-3

```

Query Match 1.2%; Score 41; DB 1; Length 1957;
Best Local Similarity 47.9%; Pred. No. 0.22;
Matches 113; Conservative 0; Mismatches 123; Indels

RESULT 10
US-08-858-003-2
; Sequence 2, Application US/08858003
; Patent No. 6060234
; GENERAL INFORMATION:
; APPLICANT: Katz, Leonard
; APPLICANT: Stassi, Diane L.
; APPLICANT: Summers Jr., Richard G.
; APPLICANT: Ruan, Xiaozan
; APPLICANT: Pereda-Lopez, Ana
; APPLICANT: Kakavas, Stephan J.
; TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
; TITLE OF INVENTION: AND RECOMBINANT METHODS FOR MAKING SAME
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Rd.
; CITY: Abbott Park
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,003
; FILING DATE: 16-MAY-1979

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-858-003-2

Query Match 1.1%; Score 40.8; DB 3; Length 1030;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 102; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 440 tccgtcagcggggaagatcgcccgagagagctgcactgcccgcgtccgagttcgacg 499
DB 203 TCACACAGCGCGCTCTGTCGCGGTGGAGACCTCCCTGTCGCGCTTCGAGGCCACG 262
QY 500 acacgttttccaagaagttcgaggtgctcttctgcccgcgtgacggtggcgcaaca 559
DB 263 GCCTCTGCCCGACTACCTCATCGGCCACTCCATCGCGGAAGTGACCGGCCACCTGG 322
QY 560 aggtctccgcggccctgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 619
DB 323 CCGGGTCTCGATCTGGCGGAGCGCTGCTGCTGCTGCCACCGCGCCGCTGATGC 382
QY 620 ggggggtccgagagccccccccc 643
DB 383 AGTCGGCGCGCGCGCGCGCA 406

RESULT 11
US-09-078-166-2
Sequence 2, Application US/09078166
Patent No. 6063561
GENERAL INFORMATION:
APPLICANT: Katz, Leonard
APPLICANT: Stassi, Diane L.
APPLICANT: Summers Jr., Richard G.
APPLICANT: Ruan, Xiaocan
APPLICANT: Pereda-Lopez, Ana
APPLICANT: Kakavas, Stephan J.
TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd.
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/078,166
FILING DATE: 16-MAY-1979
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dianne Casuto
REGISTRATION NUMBER: P-40,943
REFERENCE/DOCKET NUMBER: 4952.US.P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (847)-938-3137
TELEFAX: (847)-938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1030 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-078-166-2

Query Match 1.1%; Score 40.8; DB 3; Length 1030;
Best Local Similarity 50.0%; Pred. No. 0.18;
Matches 102; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 440 tccgtcagcggggaagatcgcccgagagagctgcactgcccgcgtccgagttcgacg 499
DB 203 TCACACAGCGCGCTCTGTCGCGGTGGAGACCTCCCTGTCGCGCTTCGAGGCCACG 262
QY 500 acacgttttccaagaagttcgaggtgctcttctgcccgcgtgacggtggcgcaaca 559
DB 263 GCCTCTGCCCGACTACCTCATCGGCCACTCCATCGCGGAAGTGACCGGCCACCTGG 322
QY 560 aggtctccgcggccctgcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 619
DB 323 CCGGGTCTCGATCTGGCGGAGCGCTGCTGCTGCTGCCACCGCGCCGCTGATGC 382
QY 620 ggggggtccgagagccccccccc 643
DB 383 AGTCGGCGCGCGCGCGCGCA 406

RESULT 12
US-08-997-467-2
Sequence 2, Application US/08997467
Patent No. 6200813
GENERAL INFORMATION:
APPLICANT: Katz, Leonard
APPLICANT: Stassi, Diane L.
APPLICANT: Summers Jr., Richard G.
APPLICANT: Ruan, Xiaocan
APPLICANT: Pereda-Lopez, Ana
APPLICANT: Kakavas, Stephan J.
TITLE OF INVENTION: NOVEL POLYKETIDE DERIVATIVES
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Rd.
CITY: Abbott Park
STATE: Illinois
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,467
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/858,003
APPLICATION NUMBER: 08/858,003
FILING DATE: 16-MAY-1997

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:07:12 ; Search time 5191.84 Seconds
(without alignments)
9241.748 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725

Perfect score: 3555
Sequence: 1 atgggaaccaataacattcac.....tccacactgtccaggccctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estin:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_htc:.*
9: gb_est1:.*
10: gb_est2:.*
11: gb_htc:.*
12: gb_gss:.*
13: em_gss_hum:.*
14: em_gss_inv:.*
15: em_gss_pin:.*
16: em_gss_vrt:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	833.6	23.4	1001	10	278359 HS278359 Hu
2	826.6	23.3	1051	10	BM479578 AGENCOURT
3	819.2	23.0	1007	10	BM016445 603641252
4	782.4	22.0	784	10	BM016445 603641252
5	741.8	20.9	835	10	BM016445 603641252
6	738.8	20.8	800	10	BM016445 603641252
7	737.4	20.7	1121	10	BM476629 AGENCOURT
8	729.8	20.5	1109	10	BM460573 AGENCOURT
9	722.2	20.3	952	10	BE299948 600944494
10	713.8	20.1	838	10	BE299948 600944494
11	705	19.8	807	10	BE299948 600944494
12	701	19.7	832	10	BM014940 603640877
13	689.8	19.4	851	9	AL570425 603640877
14	683.4	19.2	754	10	BE300005 600944594
15	682.6	19.2	892	10	BE300005 600944594
16	678.6	19.1	686	10	BE561889 601346094
17	675	19.0	1108	10	BE561889 601346094

18	671.8	18.9	978	10	BG029927
19	669.6	18.8	670	9	AL557623
c 20	669.4	18.8	857	9	AL544177
21	668.4	18.8	750	10	BI084159
22	664.6	18.7	795	10	BE560184
23	661	18.6	661	9	AL046787
24	661	18.6	948	10	BG258116
25	659.2	18.5	672	10	BE675153
26	652.8	18.4	806	9	AW043925
27	648.8	18.3	665	9	AW173375
28	647.8	18.2	812	10	BI758489
29	641	18.0	1161	10	BM466841
30	637.8	17.9	773	10	BE361312
c 31	636.4	17.9	862	10	BI910369
32	635.4	17.9	660	10	BE276812
33	632.4	17.8	761	10	BF797592
34	632	17.8	753	10	BF305442
35	630	17.7	722	10	BE560392
36	628	17.7	748	10	BG424359
37	624.4	17.6	627	9	AW772402
38	621.2	17.5	626	9	AW239183
39	612	17.2	623	10	BE884440
40	604.4	17.0	649	10	BE513375
41	597	16.8	605	9	AW732976
42	596.2	16.8	622	10	BE514108
43	592.4	16.7	774	10	BI763670
44	590.2	16.6	952	10	BG424757
45	589.4	16.6	758	10	BG166636

ALIGNMENTS

RESULT 1

278359/c

LOCUS

DEFINITION

ACCESSTION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

278359 1001 bp mRNA linear EST 28-JUL-1999
HS278359 Human fetal brain S. Meier-Ewert Homo sapiens cDNA clone
3.114 (CEPH) 5', mRNA sequence.
278359 1 GI:1495132
EST.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1001)
Meier-Ewert, S., Lebre, A.S., Holbert, S., Saada, C., Bougueleret
J., Meier-Ewert, S., Legall, I., Millasseau, P., Bui, H., Giudicelli
C., Massart, C., Guillou, S., Gervy, P., Poullier, E., Rigault, P.,
Weissenbach, J., Lennon, G., Chumakov, I., Dausset, J., Lehrach, H.,
Cohen, D. and Cann, H.M.
Survey of CAG/CTG repeats in human cDNAs representing new genes:
candidates for inherited neurological disorders
Hum. Mol. Genet. 5 (7), 1001-1009 (1996)
96414310
Contact: Neri C.
Fondation Jean Dausset - CEPH
27 Rue Juliette Dodu, 75010 Paris, France
Related sequence: 278360
5'-sequence (upper strand).
Location/Qualifiers
1. .1001
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="3.114 (CEPH)"
/clone_lib="Human fetal brain S. Meier-Ewert"
/tissue_type="brain"
/dev_stage="fetus"
/note="cDNA library of S. Meier-Ewert, Max Planck
Inst.f. Mol.Genetics, Berlin, FRG"
237 a 229 c 231 g 272 t 32 others

Query Match 23.4%; Score 833.6; DB 10; Length 1001;
Best Local Similarity 93.5%; Pred. No. 2.2e-197;
Matches 937; Conservative 0; Mismatches 56; Indels 9; Gaps 8;

QY 2146 acatctgtgagctcgaagagctgtggcaaaaaggctattcttcaacagatactgtgtt 2205
|||||
DB 1001 ACATCTCGTGTGAGCTCGAGAGCTGTGCTTAAAGGCTATTCTTCAACAGATNCTGTCTN 942
|||||
QY 2206 aqaatggagaagaaacagagctcccaagcctctgaaatgatttctgcaacagcgc 2265
|||||
DB 941 AGAATGGAGAAGAAATCAGAAAGCTCCAAAGCTCTGAAATGATTTGCTGAACAGCC 882
|||||
QY 2266 ctgaagctgattatgaagaataactccctgtcttaaaagaagtaactacagtggtgaa 2325
|||||
DB 881 CTGAAGCTCGATTATGAAGAAATNACTCCCTGTCTTAAAGAGTAACTACAGTGTGGAA 822
|||||
QY 2326 aagatgcttagcactccagaagat--caaaaaataagtttgacatggaataatgcact 2383
|||||
DB 821 AAGATGCTTAGCAGCTCCAGGAAGATCCCAANAAATTAAGTTTACATGGAANAATGCAT 762
|||||
QY 2384 cggctgttgggc-aaggtgtgcaagctatccagaggtgaaatctggaaa-ttcttagc 2441
|||||
DB 761 CGACTGTGGCTAAGTGTGCGCAGCTCATCCAGAGGTGAATCTGGAAATTTCTAGC 702
|||||
QY 2442 tgagcaa-ttccaccttaaacaccagtttcccagcaaacagcagcaaaaggtgtgcac 2500
|||||
DB 701 TGAGCAATTTCCACCTAAACACACAGTTTCCAGCANACAGCAGCCANAGATGTGCCAT 642
|||||
QY 2501 acaagaactcttaagagctgaacttcccagcagcagcagcagcagcagcagcagcagc 2560
|||||
DB 641 ACAAGAATCTTAAAGCAGCTGACTTCCAGCAGCATCGGATTCATTATGACCTTGGNC 582
|||||
QY 2561 gaaccttctcacacaccatacttctgtccagcttggagcagcagcagcagcagcagcagc 2620
|||||
DB 581 GAACCTTTCTACACACCCATCTTCTGTGCCAGCTTGGAGCAGCAGCATGCGTTT 522
|||||
QY 2621 acaaatgttgagcctactcaactctgacagcagcagcagcagcagcagcagcagcagc 2680
|||||
DB 521 ACAACATTTTGAAGGCTTACTCATCTTAGACAGGAGTGGGATATTGCCAAGGTCTCA 462
|||||
QY 2681 gctttgacagcatttgccttcttcatatgaggaagagcagcagcagcagcagcagcagc 2740
|||||
DB 461 GCTTTGACAGGCAATTTGCGGCTTCATATGATGAGGAAGAGGCGGTGAAGATGCTCA 402
|||||
QY 2741 agttctgtattgtgacatgggctcggaaacagcagcagcagcagcagcagcagcagcagc 2800
|||||
DB 401 AGTTTCTGATG-TTGACATGGGCTNNGGNACAGNATCGGCCAGACATGATTTTAC 343
|||||
QY 2801 agatccagatgtaccagctctcagaggtgttcttcagattaccacagagacctctacaac 2860
|||||
DB 342 AGANCCAGATGTACCACTCTCGGGGTTNCNTNATGATTACACAGAGACCTCAACAATC 283
|||||
QY 2861 acctgagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2919
|||||
DB 282 ACCTGAGGAGGNCAGATCGGCCAGCCTCTACGCTGCCGCCCTGGTTCCTCAGCATG 223
|||||
QY 2920 ttgctctcagctccgctgggattcgttagcagagcttcttgaatgattttcttcag 2979
|||||
DB 222 TTTGCTCAGATGTCGCTGGGATTCGTATNCAGAGNCTTTGANAATGATTTTCTTCAG 163
|||||
QY 2980 ggaacagaggtcatattaaagtggtttaaagctgttgggaagccataagcccttgatt 3039
|||||
DB 162 GGAACAGAGTCAATTTAAAGTGGCGGNAAGNCTGNTGGGAANCCATAAGCCNGTGAT 103
|||||
QY 3040 ctgacagcagaaacccatagaaa--ccatagttgactttataaaagcagcagcagcagcagc 3098
|||||
DB 102 CTCANCATGAAGAACCTGAAGNCCAGATNGACTTATGTAAAGCAGCAGCTACCCCAACCT 43
|||||
QY 3099 tggcttggcacagtgaaa-aagaccatcaatcagatttg 3139
|||||
DB 42 TGCGGTGTTNAGATGGAAGNAGACCATCAATCAGGTAGTG 1

RESULT 2
BM479578
LOCUS
DEFINITION
AGENCOURT_6464998 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577183
5', mRNA sequence.
BM479578
VERSION
BM479578.1 GI:18528620
EST.
KEYWORDS
SOURCE
Homo sapiens
ORGANISM

REFERENCE
1 (bases 1 to 1051)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM12330 row: d column: 16
High quality sequence stop: 662.

FEATURES
Location/Qualifiers

1..1051
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5577183"
/clone_lib="NIH_MGC_92"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT
296 a 278 c 234 g 243 t
ORIGIN

Query Match 23.3%; Score 826.6; DB 10; Length 1051;
Best Local Similarity 95.8%; Pred. No. 1.3e-195;
Matches 992; Conservative 0; Mismatches 34; Indels 5; Gaps 4;

QY 2451 ccaccttaaacaccagtttcccagcaaacagcagcagcagcagcagcagcagcagcagc 2510
|||||
DB 1 CCACCTTAAACACCAGTTTCCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 60
|||||
QY 2511 cttaaagcagcttccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2570
|||||
DB 61 CTTAAAGCAGCTGACTTCCAGCAGCATCGGATTCATTATGACCTTGGGGAACCTTTC 120
|||||
QY 2571 tacacaccatactctctgccagcttggcagcagcagcagcagcagcagcagcagcagcagc 2630
|||||
DB 121 TACACACCATACTTCTCTGCCAGCTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 180
|||||
QY 2631 gaagcctactcaactctctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2690
|||||
DB 181 GAAGCCTACTACTTCTAGACAGGAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGC 240
|||||
QY 2691 agcatttgccttctctatagtgaggaagcagcagcagcagcagcagcagcagcagcagcagc 2750
|||||
DB 241 AGCATTTTCTCTCTATATGAGTGAGGAGGCGTTTAAATGCTCAAGTTTCTGAT 300
|||||
QY 2751 gttgacatgggctcgggaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2810
|||||
DB 301 GTTGCATGGGCTCGGGAACAGTATCGCCAGACATGATTTATTTACAGATCCAGAT 360
|||||

QY 2811 gtaccagctctgaggttgcttcattaccacagacacgtcttacaatcacctggagga 2870
 Db 361 GTACCAGCTCTCGAGGTGCTTCATGATTACACAGAGACCTTACAATCACCTGGAGGA 420
 QY 2871 gcacgagatcgcccccagcctctacgctgcccctgggttcctcacatgtttgcctaca 2930
 Db 421 GCACGAGATCGCCCCAGCCCTCTACGCTGCCCTGGTTCCTCACCATGTTGCTCACA 480
 QY 2931 gttcccgctggagctcgtagcagagctcttgatgatgatttttcttcaggaacagaggt 2990
 Db 481 GTTCCCGCTGGGATTCGTAGCCAGAGTCTTTGATATGATGATTTTCTTCAGGGAACAGAGGT 540
 QY 2991 catatttaagtggtcttaagctctgttggaagccataagcccttgattctgcagcatga 3050
 Db 541 CATATTTAAAGTGGCTTTAAGTCGTCTGGGAAGCCATTAAGCCCTTGATCTGCGACATGA 600
 QY 3051 aaacctagaaccattagttgactcttataaaagcagcgtacccaccttggtcttggtaca 3110
 Db 601 AAACCTAGAAACCATAGTTGACTTTTATAAAAGCAGCGTACCCAACTTGGCTGTACA 660
 QY 3111 gatgaaagaccatcaatcaggtatttgaaatggagacatcgctnaaacagttacaagctta 3170
 Db 661 GATGAAAGACCATCAATCAAGTATTTGAAATGGACATCGCTAAACAGTTACAAGCTTA 720
 QY 3171 tgaagttagtaccacgtctctcaagaagaacttatcgattctctctcagtgacaa 3230
 Db 721 TGAAGTTGAGTACCAGCTCTCTCAAGAAGACTTATCATGATCTCTCTCAGTGACAA 780
 QY 3231 ccaagaatgataaattagagaacacacagcagcttaccgaaacacacaccttgacct 3290
 Db 781 CC-AAAGATGGATATATTAGAGAAACCAACAGCAGCGCTTACGAAACGAAACCTTGACCC 839
 QY 3291 cctgaacagttgcaagt-ggcataatggttaggataccaaagccttg--aggccaccattga 3347
 Db 840 TCCTTGACAGTGTGACAGTGGGCGCAATGTTAAGATCCAAAGCCCTGGAGGCCCCCATGGA 899
 QY 3348 gaagct-cctgagcagtgagagcaagctgaa 3377
 Db 900 AAAGCTCCTGNACAGGAGAGCAAGCCTAA 930

RESULT 3
 BM016445
 LOCUS 603641252F1 NIH_MGC_87 1007 bp mRNA linear EST 30-OCT-2001
 DEFINITION mRNA sequence.
 ACCESSION BM016445
 VERSION EST.
 KEYWORDS BM016445.1 Gr:16530799
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (Bases 1 to 1007)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: BCRD/BTP
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M12064 row: h column: 13
 High quality sequence stop: 822.
 FEATURES
 Location/Qualifiers
 1..1007
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5417148"

/clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 Kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 274 a 246 c 255 g 231 t 1 others
 ORIGIN
 Query Match 23.0% Score 819.2; DB 10; Length 1007;
 Best Local Similarity 95.1% Pred. No. 8.9e-194;
 Matches 910; Conservative 0; Mismatches 39; Indels 8; Gaps 6;
 QY 2532 gcagctgcgattcttattgacctggggcgaacctttctcacacccattctctctgc 2591
 Db 1 GCAGCATCGGATTCTTATTGACCTTGGCGAACCTTCTCTACACACCCATACTCTCTGC 60
 QY 2592 ccagcttgagcagagacgtatcgctttacaacattttgaggcctactcactctaga 2651
 Db 61 CCAGCTTGGAGCAGACAGCTATCGCTTTACAACATTTGAAGGCTACTCACCTCTTGA 120
 QY 2652 ccaggaagtgggattatgcccaaggtctcagctttgttagcagcattttgtctctcat 2711
 Db 121 CCAGGAAGTGGGATTTGCCAAGGCTCAGCTTTGTACAGGCANTTTGCTTCTCATAT 180
 QY 2712 gagtgaggaagagcggttttaaaatgctcaagtcttctgattgttgacatggggcgcgga 2771
 Db 181 GAGTGAGGAAGAGCGGTTTAAATGCTCAAGTTTCTGATGTTGACATGGGCTCGCGAA 240
 QY 2772 acagtatggcagacatgattattttacagatccagatgtaccagctctcagagttgct 2831
 Db 241 ACAGTATCGGCCAGACATGATTATTTTACAGATCCAGATGTACCAGCTCTCAGGTTGCT 300
 QY 2832 tcatgattaccacagagacgtctcaaatcacctggagagcacagatcgccccagcgt 2891
 Db 301 TCATGATTACACAGAGACCTTACAATCACCTGGAGGAGCAGAGATCGGCCAGCCT 360
 QY 2892 ctacgtgccccctgggtctcacaatgtttgctcagtcacagttcccgctgggattcgtagc 2951
 Db 361 CTAGCGTCCCCCTGGTTCCTCAGCATGTTTGCCTCAGATGTCCGCTGGGATTCGTAGC 420
 QY 2952 cagagctcttgatgattttttctcaggaacagagatcatatttaagtggtcttaag 3011
 Db 421 CAGAGCTCTTATGATGATTTTCTCAGGGAACAGAGTCAATTTTAAAGTGGCTTTAAG 480
 QY 3012 tctgttggaagccataagccttgattctcgcagcatgaaacccctagaaaccatagttga 3071
 Db 481 TCTGTGGGAAGCCATAGCCCTTGTATCTCAGCATGAAACCTAGAACCATAGTTGA 540
 QY 3072 cttataaaagcagcgtaccacacccctggcttgtagacagatggaaaagacatcaatca 3131
 Db 541 CTTTATAAAAGCAGCTTACCAACCTTGGCTTGGTACAGATGGAAAAGACATCAATCA 600
 QY 3132 ggtattgaaatgacatcgctaaacagttacagctctatgaagttgagtcacagctct 3191
 Db 601 GGTATCTGAAATGGACATCGCTTAACAGTTTACAAGCTTATGAAGTTGAGTACCAGCTCT 660
 QY 3192 tcaagaagaacttatcgattctcctctcagtgacaccccaagaatggataa--attag 3250
 Db 661 TCAAGAAGAACTTATCGATTCTCTCCTCTCAGTGACAACCAAGAAATGGATAACCTTAG 720
 QY 3251 agaaaacacacagcagcttacgaaaacagaccccttgacctcttgacagattga-ggtg 3309
 Db 721 AGAAAACCAACAGCAG--TTACGCAAAACAGAACCTTGGACCTCTCTTGAACAGTTGCAAGGTG 779
 QY 3310 gcaatgtagatccaaagccttgagggccaccattgaaa--gctcctgagcagtgaga 3367
 Db 780 CCAATGTTAGATCCAAAGCCTTGGAGGCCCCCATTTGAGAAAGCTCCCGGAGCAGTGA 839
 QY 3368 gcaag--ctgaagcaggccatgcttacc--ttagaactggagcggctcgccctgctgcaga 3424


```
/db_xref="taxon:9606"  
/clone="IMAGE:4851090"  
/clone_lib="NIH_MGC_48"  
/tissue_type="primary B-cells from tonsils (cell line)"  
/lab_host="DH10B (phage-resistant)"  
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;  
Site_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCAGG(G). Size-selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH_MGC Library."  
BASE COUNT 220 a 230 c 214 g 171 t  
ORIGIN  
  
Query Match 20.9%; Score 741.8; DB 10; Length 835;  
Best Local Similarity 98.6%; Pred. No. 1.9e-174;  
Matches 768; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
  
Qy 1591 gatagctccctgttagtaccttaagtaaccacgaagagccatctgtgtgtgaaag 1650  
Db 1 GATAGCTCCCTGTCTAGTACATTAAGTAACACACGACCAAGAGCCATCTGTGTGAAAG 60  
  
Qy 1651 gaggccttgccatctctgagagctcttaagctctctcggtctctcgagagacgtgcc 1710  
Db 61 GAGGCTTGGCCCATCTGTGAGAGCTCTTTAAGCTCTCGGCTCTCTCGGAGACCTGTCC 120  
  
Qy 1711 agtgactcgagagtcattctccgaagagccagctcgctgtcgccccagcagccttc 1770  
Db 121 AGTGACTCGGAGAGTCATCTCCAGAGAGGCCAGCTCCGCTGTGCCCCAGCAGCCTTC 180  
  
Qy 1771 aggaggcgagcaaacacctgtgtaattccccatcgatgccagggaacctcccaacct 1830  
Db 181 AGGAGGGGAGCAAAACACCTGAGTCACCTCCCATCGAATGCCAGGAACCTCCCAACCT 240  
  
Qy 1831 gccgggggtcccggggttcgcaagaaacattagaggtatcactcagtagcaca 1890  
Db 241 GCCCGGGGTCCCGGGGTTCGCAAGAGAACTTATGAGGTATCACTCAGTAGGAGACA 300  
  
Qy 1891 gagacgcctatgaacgaagaccttgaatccaaagcaaacacctctgtgtgattcgtg 1950  
Db 301 GAGAGCGCTCATGAACGAAGGACTTTGAATCCAAAGCAACCAACCATCTGGTGATCTGGT 360  
  
Qy 1951 gggactcctgtgaagaccggagagcatctctcgaggcagcagatattctcccgagtagcc 2010  
Db 361 GGGACTCTCTGTGAAGACCGGAGGCATCTCTGGAGGCGAGCAGATATCTCCGAGTAGCC 420  
  
Qy 2011 accccgcagaagcgtgcgattctccagcagatagaagattattcagagctggagag 2070  
Db 421 ACCCGCGAAGAGCGGTGGATCTTCAGCAGAGATATGAAGATTATTCAGAGCTGGGAGAG 480  
  
Qy 2071 ctccccacagatctctttagaaccagtgttgaaagatggccctttggccccccacca 2130  
Db 481 CTTCCTCCACCATCTCTTTAGAACACAGTTGTGAAGATGGCCCTTTGGCCCCCACCACCA 540  
  
Qy 2131 gaggaaagaagaagacatctcgtgagctcgcagagctgtggcaaaaggctattcttcaa 2190  
Db 541 GAGGAAAGAAAGAGACATCTCGTGTGAGCTCGGAGAGCTGTGGCAAAAGGCTATTCTTTCAA 600  
  
Qy 2191 cagatactcgtctagaaatggagaagaaatacagaagctccaagcctctgaaatgat 2250  
Db 601 CAGATATCTGCTGTAGAAATGGAGAAGAAATCAGAAAGCTCCAAGGCTCTGAACATGAT 660  
  
Qy 2251 ttgctgaacagcgcctgaagctcgattgaagaataattcctcctgtctttaaagaagta 2310  
Db 661 TTGCTGAACAGCGCCTGAAGCTCGATTATGACGAATTAATTACTCCTGCTTTAAGAAGATA 720  
  
Qy 2311 actacagtg- gggaaagatgctta- gcactccaggaagaatcaaaaattgaattg 2365  
Db 721 ACTACAGTGTGGGAACAGATGCTTAGGCACTCCAGGAAGATCAACAATCAGGTTGC 777
```

RESULT 6
BG746377
LOCUS
DEFINITION
602703644F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4856891 5',
mRNA sequence.
BG746377
ACCESSION
BG746377.1 GI:14057030
VERSION
BG746377.1
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 800)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI708 row: h column: 12
High quality sequence stop: 762.
FEATURES
source
1..800
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4856891"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
BASE COUNT 213 a 218 c 203 g 166 t
ORIGIN

Query Match 20.8%; Score 738.8; DB 10; Length 800;
Best Local Similarity 98.6%; Pred. No. 1.1e-173;
Matches 787; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

Qy 1502 agagattcttaacagagctcttgaagaagtattttgtcccggtgtaataaagccagagcc 1561
Db 2 AGAGATCTTTAAACAGAGCTCTTTAGAAAGATATTTGTCCGGGTAAATAAGCCAGAGGCC 61

Qy 1562 tgaggaaactccatccatcagtggtgagtagctccctctgtctagtagacattagtaaca 1621
Db 62 TGCAGGAACACTCCATCAGTGTGATCTGGATCTGCTCTCTAGTACATTAAGTAAACA 121

Qy 1622 ccagcaagagcattctgtgtgaaagagggccttgccatctctgagagctccctta 1681
Db 122 CCAGCAAGAGCCATCTGTGTGTGAAAGAGGCGCTTGCCCATCTCTGAGAGCTCCTTTA 181

Qy 1682 agctcctcgctctcgcggagacgtctccagtgactcgagtagctatctcccagaagagc 1741
Db 182 AGCTCTCGCTCTCTCGGAGGACCTGCTCCAGTGACTCGGAGAGTCATCTCCAGAAGAGC 241

Qy 1742 cagctcgcgtctgcgccccagcagccttcaggagcgagaaacacctgagtcacttc 1801
Db 242 CAGCTCCGCTGTGCGCCCGCAGCAGGCCCTTCAGGAGCGGCAACACCCCTGAGTCACCTTC 301


```

DEFINITION  AGENCOURT_6421241 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5531908
5', mRNA sequence.
ACCESSION   BM460573
VERSION     BM460573.1 GI:18509613
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1109)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12214 row: f column: 05
            High quality sequence stop: 630.

FEATURES             source
     source
     1..1109
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="IMAGE:5531908"
         /clone_lib="NIH_MGC_71"
         /tissue_type="leiomyosarcoma"
         /lab_host="DH10B (phage-resistant)"
         /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
            Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 2.1 kb."
BASE COUNT  301 a 311 c 258 g 239 t
ORIGIN
|||||
Query Match      20.5%; Score 729.8; DB 10; Length 1109;
Best Local Similarity 97.5%; Pred. No. 2.2e-171;
Matches 773; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 1340 tgattattctttctgagatgtttatatgaagagaacaaagaaacacatccatttg 1399
Db 13  tGATTATTCTTTCTGAGATGTTTATGAGAGAAACAGAAAGACATCCCATATTG 72

QY 1400 gggagatgaagcagacatcgagatggcagcagagaattatggagtgaaattaccacca 1459
Db 73  GGGAGATGAGCAGACATCGCAGATGGCAGCAGAGAAATTTGGAAGTGAATACCAACCA 132

QY 1460 gtgccactcgattttaggttagatgtctgtaaaacaaagcaagagatctttaacagagt 1519
Db 133 GTGCCACTCGATTAGGCTAGATGCTGTAAGAAACAAAGCAAGAGATCTTTAACAGAGT 192

QY 1520 ctttgaagaattttgtcccggtgtaataaagcagagcgctgcagaaacactccatca 1579
Db 193 CTTTGAAGAATTTTGTCCCGGGGTAAATAAGCCAGAGGCGCTGCAGGAACACTCCATCA 252

QY 1580 gtgtgagctgtagtgcctctgtctagatattgaataacacacagcaagagcattctg 1639
Db 253 GTGTGATCTGATAGTCCCTGCTCTAGTACATTAAATTAACACAGCAAGAGCCATCTG 312

QY 1640 tgtgtgaagaagggccttgcccatctcttgagagctcttttaagctctctgctctcgtg 1699
Db 313 TGTGTGAAGAAGGAGGCGCTTGCCCATCTCTGAGAGCTCCTTTAAGCTCTCGGCTCTCGG 372

QY 1700 aggaactgtccagtgaactcgagagatcattctccagaagaccagctcgctgtgccccc 1759
Db 373 AGGAGCTGTCCAGTACATCGGAGATCATCTCCCAAGAGAGCCAGCTCCGCTGTGCCCCC 432

QY 1760 agcaggcccttcaggagcgagcaaacacctgagtcacttccccatcgaaatgccaggaa 1819
|||||

```

```

Db 433 AGCAGGCTTCAGGAGCGAGCAAAACACCCCTGAGTCACCTCCCATCGAATGCCAGAAC 492
QY 1820 ctccaaacctgcccgggggtcccccgggggtttccaaagaaacttatgaggtatcaact 1879
Db 493 CTCCCAAACCTGCCCGGGGTCCCGGGGGTTTCGCAAGAGAACTTATGAGGTATCACT 552
QY 1880 cagtgagcacagagacgctcatgaacaaagacatttgaatccaaagcaaacacattctg 1939
Db 553 CAGTGAGCACAGAGACGCTCATGAACGAAGAGACTTTGAATCCAAAGCAAAACCATCTTG 612
QY 1940 gtgattctgtgtgggactcctgtgaagaccggagggagcattcctctggagcagcatattcc 1999
Db 613 GTGATTCTGTGGGACTCCTGTGAAGACCCCGGAGCATTTCTCTGGAGCAGCATATTCC 672
QY 2000 tccgagtagccaccgcccgagagggcgctgctctccagcagatatgaagattattcag 2059
Db 673 TCCGAGTAGCCACCCCGCAGAAAGCGGGCGATTTCTTCAGCAGATATGAGATTTATCCG 732
QY 2060 agct-ggagagactccccccacgac-tcctttagaacacagtttgt-gaagatgggccc 2116
Db 733 AGCTGGGAGAGACTTCCCCCCCCCAACTTCCCTTAGAACAGTTTGTGGAGATGGGGCC 792
QY 2117 ttggccccccacc 2129
Db 793 TTTGGCGCCTCC 805

RESULT 9
BE299948
LOCUS     BE299948      600944494F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960782 5',
DEFINITION mRNA sequence.
ACCESSION BE299948
VERSION   BE299948.1 GI:9183696
KEYWORDS EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 952)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: L1CM54 row: k column: 23
            High quality sequence stop: 709.

FEATURES             source
     source
     1..952
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /clone="IMAGE:2960782"
         /clone_lib="NIH_MGC_17"
         /tissue_type="rhabdomyosarcoma"
         /lab_host="DH10B (phage-resistant)"
         /note="Organ: muscle; Vector: pOT87; Site_1: EcoRI;
            Site_2: XhoI; cDNA made by oligo-dT priming.
            Directionally cloned into EcoRI/XhoI sites using the
            following 5' adaptor: GGCAGAG(G). Size selected by
            for average insert size 1.8kb. Library constructed by
            Ling Hong in the laboratory of Gerald M. Rubin (University
            of California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT  249 a 240 c 250 g 212 t 1 others
ORIGIN
|||||

```

Query Match 20.3% Score 722.2; DB 10; Length 952;
Best Local Similarity 97.2%; Pred. No. 1.7e-169;
Matches 766; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

QY 1879 tcagtgcagacagagcgcctcatgaacgaagacttgaatccaaagcaaacacattctt 1938
Db 1 TCAGTGAGCACAGACAGCGCCTCATGAAGCAAGAGGACTTTGAATCCAAAGCAACACATCTT 60

QY 1939 ggtgattctggtgggactcctgtgaagacccggaggcattcctggaggagcagatattc 1998
Db 61 GGTGATTCTGTGGGACTCCTGTGAAGACCCGGAGGACTTCCCTGGAGGAGCAGATATTC 120

QY 1999 ctccgagtaccaccgccgagaaagcggtgcattctccagcagatatgaagattattca 2058
Db 121 CTCCGAGTAGCCACCCGAGAGAGGGGTGCGATCTTCAGCAGATATGAGATTTATTC 180

QY 2059 gagctgggagagcttcccccaagatctcctttagaaccagttgtgaagatgggcccctt 2118
Db 181 GAGCTGGAGAGCTTCCCCCAGCATCTCTTTAGAACACAGTTTGTGAAGATGGCCCTTT 240

QY 2119 ggcctcccccacagagaaagaaagacatctcgtgagctcagagagctgtggcaaaag 2178
Db 241 GGCCTCCCCACAGAGAAAGAAAGACATCTCGTGAAGTCCGAGAGCTGTGGCAAAAG 300

QY 2179 gctattctcaacagatactgctgttagaatgagaagaaatcagaagctcccaagcc 2238
Db 301 GCTATCTTCAACAGATAGTCTGCTTAGAATGGAGAGGAAATCAGAGCTCCAGGCC 360

QY 2239 tctgaataatgattgtgaacagcgcctgaagctcgattatgaagaaattactccctgt 2298
Db 361 TCTGAAATGATTGTGTAACAAGCGCCTGAAGCTCGATTATGAAGAAATTAATCTCCTGT 420

QY 2299 cttaagaagtaactacagtggtggaaagatgcttagcaactccaggaagatcaaaaatt 2358
Db 421 CTTAAGAAGTAACAGTGTGGGAAAGATGCTTAGCACTCCAGGAAGATCAAAAT 480

QY 2359 aagttgacatgaaataatcgactcgctgttgggcaaggtgcccagctcatcacga 2418
Db 481 AAGTTTGACATGAAATAATGCCTCGGCTG-TGGCAAGGTGTGCACGCTCATCCGA 539

QY 2419 ggtgaatctggaaattctagctgagcaattccacccttaaacacagtttcccgacaa 2478
Db 540 GGTGAATCTGGAATTTCTAGCTGAGCAATCCACCTTAACACACAGTTTCCAGCAAA 599

QY 2479 cagcagcaaaagatgtgccatacaagaactcttaagcagctgacttcccgagcagcat 2538
Db 600 CAGCAGCAAAAGGATGTCCATCAAGAAGTCTTAAAGCAGCTGACTTCCAGCAGCAT 659

QY 2539 gcgattctta-ttgaccttggggaaaccttccctacacacccactacttctctgccagct 2597
Db 660 GCGATTCTTATTGACCTGNGGCGAACCTTTCCTACACACCCATCTCTCTGCCAG-T 718

QY 2598 tggagcagcagcagctatcgctttacaacattttgaagcgctactcacttctagaccagga 2657
Db 719 TGGAGCAGCAGCAGCTATCGCTTTACAACATTTGTGAGGGCTACTCCACTTCTAGACAGGA 778

QY 2658 agtgggat 2665
Db 779 GTGGGTAT 786

RESULT 10
BI907513
LOCUS 838 bp mRNA linear EST 16-OCT-2001
DEFINITION 603065640F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5214945 5',
mRNA sequence.
ACCESSION BI907513
VERSION BI907513.1 GI:16170347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 838)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>;
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11539 row: O column: 10
High quality sequence start: 4
High quality sequence stop: 780.
Location/Qualifiers
1. 838
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5214945"
/clone_lib="NIH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."
BASE COUNT 237 a 199 c 191 g 211 t
ORIGIN

Query Match 20.1%; Score 713.8; DB 10; Length 838;
Best Local Similarity 97.0%; Pred. No. 2e-167;
Matches 812; Conservative 0; Mismatches 17; Indels 8; Gaps 8;

QY 2226 gaagctccaaagcctcgaataatgattgtgcaacagcgcctgaagctcatttatgaaga 2285
Db 9 GAAGCTCCAAAGCCTCTGAAATGATTGCTGGAACAAGCGCCTGAAGCTCGATTATGAAGA 68

QY 2286 aattactccctctctaaagaagtaactacagtggtgggaaagatgcttagcactccag 2345
Db 69 AATTACTCCCTGCTTTAAAGAAGTAAGTACTAGTGTGGGAAAGATGCTTAGCAGCTCCAGG 128

QY 2346 aagatcaaaaattaaagtttgacatggaataaaatgcactcggctgtgggcaaggtgtgcc 2405
Db 129 AAGATCAAAAATTAAGTTTGACATGGAAAAATGCACCTCGCTGTGGCAAGGTGTGCC 189

QY 2406 acgtcatcaccgagggtgaataatcggaaattctagctgagcaattccacatttaaacacca 2465
Db 189 ACGTCTATCACCGAGGTGAAATCTGGAATTTCTAGCTGAGCAATTCACCTTTAAACACCA 248

QY 2466 gttccccagcaaacagcagcgaagatgtgccatacaagaactcttaagcagctgac 2525
Db 249 GTTTCCAGCAAAACAGCAGCCAAAGGATGTGCCATACAAAGACTCTTAAGCAGCTGAC 308

QY 2526 ttcccagcagcagctgattctattgaccttggggcaaccttctctacacacccatactt 2595
Db 309 TTCCCAGCAGCATGCGATTCTTA-TGACCTTGGGGAACCTTTCTCTACACACCATACTT 367

QY 2586 ctctgcccagcttggagcagcagcagctatcgctttacaacattttgagcagcctactcact 2645
Db 368 CTCTGCCAGCTTGGAGCAGCAGCAGCTATCGCTTTTACAACA-TTTGAAGGCCCTACTCACT 426

QY 2646 tctagaccaggaagtggtgatattgccaaggtctcagctttgttagcagcagctatttgcctct 2705
Db 427 TCTAGACCAGGAAGTGGGATA-TGCCAAGGTCTCAGC-TTGTAGCAGGCA-TTGTGCTTCT 483

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: c9abps-remail.nih.gov
 Tissue Procurement: jCSTD/Drp
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLML2063 row: j column: 19
 High quality sequence stop: 811.
 Location/Qualifiers
 1. .832
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5416818"
 /clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="PH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 231 a 205 c 186 g 208 t 2 others
 ORIGIN

Query Match 19.7%; Score 701; DB 10; Length 832;
 Best Local Similarity 97.5%; Pred. No. 3.le-164;
 Matches 785; Conservative 0; Mismatches 12; Indels 8; Gaps 7;

QY 2405 cagctcatcaccaggtgaaatctgaaattctagctgagcaattccaccttaaacacc 2464
 |||||
 Db 29 CAGCTCATCACCAGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCACCTTTAAACACC 88
 |||||

QY 2465 agttcccccagcaacagcagccaaagatgctgcatacaaaagactcttaaaagcagtga 2524
 |||||
 Db 89 AGTTTCCACGAAACAGCAGCCAAAGGATGTCATCAAAAGAACTCTTAAAGCAGCTGA 148
 |||||

QY 2525 ctcccagcagcagct 2584
 |||||

Db 149 CTTCCAGCAGCATGCGATTTCTTA-TGACCTTGGCGGAACCTTTCTTACACACCCATACT 207
 |||||

QY 2585 tctctgccagctgagcagcagcagctatcgctttacacatttgaaagcctactcac 2644
 |||||

Db 208 TCTCTGCCAGCTTGGAGCAGCAGCTATCGCTTTACACATTTTGAAGCCCTACTCAC 267
 |||||

QY 2645 tcttagaccaggaagtggatattgccaaggtctcagctttgttagcaggcatctttcttc 2704
 |||||

Db 268 TTCTAGACCAGGAAGTGGGATATTGCCAAGGTCTCAGCTTTGTAGCAGGCATTTGTGCTC 327
 |||||

QY 2705 tcatatgagtgaggaagcggcttttaaatgctcaagttctgagctttgacatggggc 2764
 |||||

Db 328 TTCAATAGCTGAGGAAGAGCGCTTTAAATGTCTCAAGTTTCTGATGTGTGACATGGGCG 387
 |||||

QY 2765 tgcggaacagctatcgccagacatgattttacagatccagatgacagctctcga 2824
 |||||

Db 388 TCGGGAACAGATATCGCCAGACATGATTTTACAGATCCAGATGACAGCTCTCGA 447
 |||||

QY 2825 gggtgtctcatgattaccacagagacctctacaatcaccttggaggagcagagatggcc 2884
 |||||

Db 448 GGTGTCTCATGATTACACAGACCTCTACAATCACCTTGGAGGAGCAGAGATCGGCC 507
 |||||

QY 2885 ccagctctcagctgccccctgttctcaccatgtttgctcaccagcttccccctgggat 2944
 |||||

Db 508 CCAGCCTCTACGCTGCCCGCTGGTCTCTCACCATGTTTGTGCTCAGATTCCTCCGCTGGGAT 567
 |||||

QY 2945 tcgtagcagagcttttgatgattttcttcagggaacagagctcatatttaa-agt 3003
 |||||

Db 568 TCGTAGCCAGAGCTTTGATATGATTTTCTTCAGGGAACAGAGGTCATATTTAACAGTG 627
 |||||

QY 3004 gctttaagtctgttgggaagccataagcccttgattctgcagcatgaaaaacctagaaacc 3063
 |||||

Db 628 GCTTTAAGTCTGNTGGGAAGCCATAAGCCCTTGATTCTGCAGCATGAAAACCTAGAAACC 687
 |||||

QY 3064 atagttgacttat-aaaaagcagctaccacaccttggc-ttggtagaat-ggaaaag 3120
 |||||

Db 688 ATAGTTGATTTTATAAAAAAGCAGCTTAACCAAGCTTGGCTTTGGTACAGATGGGAAG 747
 |||||

QY 3121 accatacaatcaggtatttgaatggacatcgctaaa--cagttacacagcttatgaagttg 3178
 |||||

Db 748 ACCATCAATCAGGTATTGGAATGGACATCGTAAACACAGTTACAAAGCTTATGAAGTTG 807
 |||||

QY 3179 agtacc-acgtccttcaagaagaac 3202
 |||||

Db 808 AGTACCAACGCTCTTCAAGAAGAAC 832
 |||||

RESULT 13

AL570425

LOCUS AL570425 LTI_NFL006.PL2 Homo sapiens cDNA clone CSODI020YF02 5
 DEFINITION prime, mRNA sequence.

ACCESSION

AL570425
 AL570425.1 GI:12926720

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 851)

AUTHORS

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source
 1. .851
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CSODI020YF02"
 /clone_lib="LTI_NFL006.PL2"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site: 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com>"

BASE COUNT 240 a 214 c 220 g 174 t 3 others
 ORIGIN

Query Match 19.4%; Score 689.8; DB 9; Length 851;
 Best Local Similarity 98.9%; Pred. No. 2e-161;
 Matches 714; Conservative 1; Mismatches 4; Indels 3; Gaps 2;

QY 2836 gattaccacagagacacctctacaat--cacctggagagacagagatcgccacagcctct 2893
 |||||

Db 1 GATTACCACAGAGACCTCTCAANTCACCTTGGAGGAGCAGAGATCGGCCCGCCGCTCT 60
 |||||

QY 2894 acgtgccccctgttctcaccatgttgcctcacagttccgctgggattcgtagcca 2953
 |||||

Db 61 ACCTGCCCCCTGGTTCCTCACCATGTTTCCCTCACAGTTCCCGCTGGGATTCGTAGCCA 120
 |||||

QY 2954 gactcttgatagattttcttcagggaacagaggtcatatttaaagtggctttaagtc 3013
 |||||

Db 121 GAGCTTTGATGATGATTTTCTTCAGGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTC 180
 |||||

```
QY 3014 ttgttggaagcacaagcccttgattctgcagcatgaaaaacccatagaacacattgaact 3073
Db 181 YGTTGGGAGGCATTAAGCCCTTGATCTGACAGTGAAGAACTTAGAAACCATAGTTGACT 240
QY 3074 ttataaagcagcgtaccccaaccttggttggtacagatggaagaccatcaatcagg 3133
Db 241 TTATAAAAAGCAGCGTACCCCAACCTTGCTGGTACAGATGGAAGAACCATCAATCAGG 300
QY 3134 tatttgaatggacatgctaaacagtttacaagcttataagcttataagcttgagttaccacgtcttc 3193
Db 301 TATTGGAATGGACATCGCTAAACAGTTTACAAGCTTATGAAGTTTGAAGTACCAACGCTCCCTTC 360
QY 3194 aagaagaacttatgattccctctctctcagtgagacaaacaaagaatgataaattaga 3253
Db 361 AAGAAGAATTTATCGATTCTCTCTCTCAGTGACAAACCAAGAAATGGATTAATTAGAGA 420
QY 3254 aaaccaacagcagcttaccgaaacagaaaccttgacctcttgaaacagttgagtgagga 3313
Db 421 AAACCAACAGCAGCTTACGCAAAACAGAAACCTTGACCTCTTGAACAGTTGCAGGTGGCA 480
QY 3314 atgtaggatacaaaagccttgagcccaacattgagaagctcctcctgagcagtgaggaagc 3373
Db 481 ATGGTAGGATCAAAAGCCTTGAGGCCACCATTTGAGAGCTCCTGAGCAGTGAGAGCAAGC 540
QY 3374 tgaagcagggcattgcttaccttagaaactgagcggtcgccctgctgcagacggtgagg 3433
Db 541 TGAAGCAGGCCATGCTTACCTTAGAACTGAGCGGCTGGCCCTCTGCTGCAGACGGTGGAGG 600
QY 3434 agctgcggcgagcagcagccagcagcgccggagcctgagtgagtcagcagcagccagc 3493
Db 601 AGCTGCGCGCGGA-SGCAGAGCCAGCGAGCGGAGCCTGAGTGACGCGAGCCCGAGC 659
QY 3494 ccagggcgactgacagctctcaggagagattgcaacacacatccacactgtccagggc 3553
Db 660 CCAGCGGCGACTGACGCTGCTGAGGAGAGATTGCACACCATCCACACTGTCCAGGCC 719
QY 3554 tt 3555
Db 720 TT 721

RESULT 14
BE300005
LOCUS 60944594F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:2960783 5',
DEFINITION mRNA sequence.
ACCESSION BE300005
VERSION BE300005.1 GI:9183753
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1C54 row: k column: 24
High quality sequence start: 2
High quality sequence stop: 753.
Location/Qualifiers
1. .754
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2960783"
```

```
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 227 a 186 c 177 g 164 t
ORIGIN

Query Match 19.2%; Score 683.4; DB 10; Length 754;
Best Local Similarity 99.1%; Pred. No. 7.5e-160;
Matches 750; Conservative 0; Mismatches 1; Indels 6; Gaps 6;

QY 1878 ctcagtgagcacagagagcgcctcatgaagaaagactttgaatccaaagcaaacacatct 1937
Db 1 CTCAGTGAGCACAGAGAGCGCCTCATGAACGAAGAGACTTTGAATCCAAAGCAAAACCATC- 59
QY 1938 tggtagctctggtggactcctgtgaagaccggaggcattctcctggagcgagcagatatt 1997
Db 60 TGGTGATTCCTGGTGGGACTCCTGTGAAGACCCGGGAGGCAATTCCTGGAGGCAGAGATATT 119
QY 1998 cctccgagtagccaccgccgagaaggcgtgcgattcttccagcagatatgaagattattc 2057
Db 120 CCTCCGAGTAGCCACCCCGCAGAGAGGGGTGCGATTCTTCCAGCAGATATGAAGATTATTC 179
QY 2058 agaactgggagagcttccccacagatctcctttagaaccagttgtgaagatgggccctt 2117
Db 180 AGAGCTGGGAGAGCTTCCCCCAGCATCTCTTTAGAACCCAG-TTGTGAAGATGGGCCCT- 237
QY 2118 tggcccccaccagaggaagaaagacatctctgagctccgagctggtggcaaaa 2177
Db 238 TGGCCCCCACCAGGAGAAAGAAAGGACATCTCGTGAGCTCCGAGAGCTGTGSCAAA 297
QY 2178 ggcattcttcaacagatactgctgcttagaattggagaagaaatacagaagctccaaagc 2237
Db 298 GGCATTATTCACACAGATACTGCTGCTTAGAATGAGAGAAATCAGAAAGCTCCAAGC 357
QY 2238 ctctgaaaaattgtctgaacagccctgaagctgattatgaagaaattactccctg 2297
Db 358 CTCTGAAAAATGA-TTGCTGAACAAAGCCCTGAAAGCTCGATTATGAAGAAATTTACTCCCTG 416
QY 2298 tcttaagaagtaactacagtggtggaaaaagatgcttagcactccaggaagatcaaaaaat 2357
Db 417 TCTTAAAGAAAGTAATACTACAGTGTGGGAAAAGATGCTTAGCAGCTCCAGAGATCAAAAAT 476
QY 2358 taegttagacatggaaaaataatgactcggctgtgtggcaaggtgtgccagctatcacccg 2417
Db 477 TAAGTTTGACATGGAATAAATGCACTCGCTGTGTGGCAAGGTGTGCCACGCTCATCACCG 536
QY 2418 aggtgaaatctggaaattcttagctgagcaattccacttaaacaccagtttccacagcaa 2477
Db 537 AGGTGAAATCTGGAAATTTCTAGCTGAGCAATTCACCTTTAAACACACAGTTTCCACGCA 596
QY 2478 acagacccaaagagtgatgcatacaaaagaaactttaagcagctgacttccccagcagca 2537
Db 597 ACAGCAGCCAAAGGATGTGCCATACAAAGAACTCTTAAAGCAGCTGACTTCCAGCAGCA 656
QY 2538 tgcattcttattgacottggcgaaaccttctacacacccatctctctctgcccagct 2597
Db 657 TGCATTCTTTA-TGACCTTGGCGAAGCTTTCTTACACACCACTACTTCTTGTGCCAGCT 715
QY 2598 tggagcaggacagctatcgctttacaacatttgaag 2634
Db 716 TGGAGCAGGACAGCTATCGC-TTAAACATTTTGAGG 751

RESULT 15

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:27:02 ; Search time 706.85 Seconds
(without alignments)
8634.973 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725

Perfect score: 3555
Sequence: 1 atgggaacaaataacattccac.....tccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3555	100.0	3983	21	AAZ50906 Human TBC-1 cDNA f
2	3555	100.0	3988	21	AAZ50907 Human TBC-1 cDNA f
3	3026.4	85.1	3326	22	AAI57854 Human polynucleoti
4	2940.2	82.7	3727	21	AACT7404 Human ORFX ORF2959
5	2833.8	79.7	3281	22	AAI59640 Human polynucleoti
6	2510.4	70.6	4039	19	AAV05886 Human Tbc-1 gene.
7	2112	59.4	2362	22	AAH18401 Human cDNA sequenc
8	1000.8	28.2	1422	22	AAI54787 cDNA encoding nove
9	999.4	28.1	2989	22	AAK94796 Human full-length

10	943.2	26.5	1269	22	AAI50906	cDNA encoding nove
11	785	22.1	883	22	AAH04450	Human cDNA clone (
12	574.2	16.2	849	22	AAK92006	Human cDNA 5'-end
13	574.2	16.2	849	22	AAK93666	Human cDNA-clone r
14	422.6	11.9	17590	21	AAZ50904	Human TBC-1 partia
15	362	10.2	696	20	ABLI5229	Human gene express
16	306.2	8.6	5472	23	ABLO2013	Drosophila melanog
17	284.4	8.0	99960	21	AAZ50905	Human TBC-1 partia
18	280	7.9	458	20	AAZ50905	Human secreted pro
19	245.6	6.9	8105	23	ABLO2012	Drosophila melanog
20	238.4	6.4	555	22	AAH09958	Human cDNA clone (
21	132.8	3.7	3691	22	AAH58616	Human RECAP polynu
22	132.8	3.7	4827	22	AAI59592	Human polynucleoti
23	132.8	3.7	5294	22	AAI57806	Human polynucleoti
24	128.8	3.6	3585	23	ABLI2337	Drosophila melanog
25	123.8	3.5	154	23	AAI58636	cDNA #1312 encodin
26	119.4	3.4	2626	23	ABLI18133	Drosophila melanog
27	119.4	3.4	7829	23	ABLI18133	Drosophila melanog
28	108	3.0	2424	23	AAI571793	DNA encoding novel
29	83.8	2.4	1601	21	AAI33401	Arabidopsis thalia
30	79.6	2.2	6765	23	ABLI12336	Drosophila melanog
31	78.4	2.2	300	20	AAI13067	Human gene express
32	75.2	2.1	1492	21	AAI38750	Arabidopsis thalia
33	73.8	2.1	1140	21	AAZ97020	Human secreted pro
34	68.4	1.9	3342	22	AAI32789	Human secreted pro
35	68	1.9	2492	21	AAI78199	Human cancer assoc
36	65.2	1.8	1831	22	AAH64964	Human secreted pro
37	65	1.8	1117	22	AAI61302	Human polynucleoti
38	65	1.8	2035	22	AAI59600	Human cell cycle a
39	65	1.8	2482	22	AAI59516	Human polynucleoti
40	65	1.8	2666	22	AAH17866	Human cDNA sequenc
41	62.6	1.8	2531	23	ABLO7359	Drosophila melanog
42	62.6	1.8	4661	23	ABLO7358	Drosophila melanog
43	61.6	1.7	2801	22	AAI32643	Human cDNA encodin
44	61.6	1.7	2803	22	AAI27080	cDNA encoding nove
45	61.6	1.7	2903	22	AAI51641	Human polynucleoti

ALIGNMENTS

RESULT 1

AAZ50906
ID AAZ50906 standard; cDNA; 3983 BP.

XX AAZ50906;

XX 31-MAY-2000 (first entry)

XX Human TBC-1 cDNA from first transcript.

XX TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;
XX Single nucleotide polymorphism; tissue differentiation; prostate cancer;
XX linkage analysis; genetic map; detection; diagnosis; genotyping;
XX transgenic animal; screening; alternative splicing; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..170

FT /*tag= a

FT /*note= "Spans through exon 1 and part of exon 2"

FT CDS 171..3677

FT /*tag= b

FT /*product= "Human TBC-1 protein"

FT /*note= "First transcription product"

FT 3'UTR 3726..3983

FT /*tag= c

FT polyA_signal 3942..3947

FT /*tag= d

XX WO200008209-A2.

XX

KW Single nucleotide polymorphism; tissue differentiation; prostate cancer;
 KW linkage analysis; genetic map; detection; diagnosis; genotyping;
 KW transgenic animal; screening; alternative splicing; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX 5'UTR 1..175

XX misc_feature 1..438

XX /tag- a

XX /tag- b

XX /note- "Homologous to 5' EST obtained from human

XX pancreas cDNA library"

XX CDS 176..3682

XX /tag- c

XX /product- "Human TBC-1 protein"

XX /note- "Second transcription product"

XX 3'UTR 3731..3988

XX /tag- d

XX polyA_signal 3947..3952

XX /tag- e

XX WO200008209-A2.

PN 17-FEB-2000.

XX 06-AUG-1999;

XX 99WO-1B01444.

XX 07-AUG-1998;

XX 98US-0095653.

XX (GEST) GENSET.

PA Blumenfeld M, Bougueret L, Chumakov I;

PI WPI; 2000-205736/18.

XX P-PSDB; AAY45096.

DR New isolated human TBC-1 nucleic acids, useful for developing products

XX for the diagnosis and treatment of disorders involving cell

XX proliferation, particularly prostate cancer

XX Claim 4; Page 156-161; 166pp; English.

XX The present sequence is the cDNA encoding the second transcription

CC product of human TBC-1 protein, comprising exons 1b1s to 2 and A-L.

CC Alternative splicing events result in two mRNA molecules, due to

CC splicing at two distinct first exons, exon1 and exon 1b1s. TBC-1 gene

CC is mapped to a candidate region of prostate cancer on chromosome 4.

CC TBC-1 gene is involved in the regulation of cell cycle and tissue

CC differentiation in mammals. An alteration of TBC-1 sequence may be

CC associated with a pathological condition, resulting in abnormal cell

CC proliferation leading to cancer, e.g. prostate cancer. The biallelic

CC markers can be used for generation of genetic maps, linkage analysis and

CC association studies. TBC-1 sequence can be used for detection,

CC diagnosis, genotyping, production of transgenic animals and screening

CC of compounds for use in therapy.

XX Sequence 3988 BP; 1107 A; 983 C; 1002 G; 896 T; 0 other;

Query Match 100.0% Score 3555; DB 21; Length 3988;
 Best Local Similarity 100.0% Pred. No. 0;
 Matches 3555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggaaccataacattcacagaagaacattcttcttaacagggtctcgttgat 60

DB 176 atggaaccataacattcacagaagaacattcttcttaacagggtctcgttgat 235

QY 61 ttggcctcagctgtgggtccctcgtgtgattccctgacacacatgcccattgctg 120

DB 236 ttggcctcagctgtgggtccctcgtgtgattccctgacacacatgcccattgctg 295

QY 121 ccttggtgtggtgagtgaggaactcagcagcagcagtcaccacagaagaacctgta 180

DB 296 ccttggtgtggtgagtgaggaactcagcagcagtcaccacagaagaacctgta 355

QY 181 accaagcaagtcgggtttgctttccacctctggtgagatgtgaacctgagccagg 240

DB 356 accaagcaagtcgggtttgctttccacctctggtgagatgtgaacctgagccagg 415

QY 241 agaagtcacagtcgggtccctcgtgattccagcatctttgagtcgaagcctcagct 300

DB 416 agaagtcacagtcgggtccctcgtgattccagcatctttgagtcgaagcctcagct 475

QY 301 gttcacaactgattcacacagtcaccccaagttactttgctgtgattaaggaa 360

DB 476 gttcacaactgattcacacagtcaccccaagttactttgctgtgattaaggaa 535

QY 361 gacgtgtccacggcagagtcattgtgtttcaaaagcagatcatacaaaagt 420

DB 536 gacgtgtccacggcagagtcattgtgtttcaaaagcagatcatacaaaagt 595

QY 421 cctgagatcatcagtcctccatccgtcagggggaagatcccgccagagagctgcac 480

DB 596 cctgagatcatcagtcctccatccgtcagggggaagatcccgccagagagctgcac 655

QY 481 tgccgtccagtcgagtcgacacagcttttccaaagattcgaggtgcttcttcggtccgc 540

DB 656 tgccgtccagtcgagtcgacacagcttttccaaagattcgaggtgcttcttcggtccgc 715

QY 541 gtgacgtgtggtgcacaaagtcctcggccctgacagagtcgacagagtcgagagttc 600

DB 716 gtgacgtgtggtgcacaaagtcctcggccctgacagagtcgacagagtcgagagttc 775

QY 601 aatcacgtcagcggcagccgggtccgagagcccccagcccccaccccccagcccg 660

DB 776 aatcacgtcagcggcagccgggtccgagagcccccagcccccaccccccagcccg 835

QY 661 ccacagggcagccagagcctgtgcgagcccatcgcaagtccttctccagccggc 720

DB 836 ccacagggcagccagagcctgtgcgagcccatcgcaagtccttctccagccggc 895

QY 721 ctgctcgtcgtggttttaggaagtcgaggtgaggggtcccgagcagcagcgttc 780

DB 896 ctgctcgtcgtggttttaggaagtcgaggtgaggggtcccgagcagcagcgttc 955

QY 781 ttacgtccttcgagagagcagcattgagaccctcattagcggacacataatgtg 840

DB 956 ttacgtccttcgagagagcagcattgagaccctcattagcggacacataatgtg 1015

QY 841 cagccacagatcagaggaataatcgactgtcttcacgattgcccagtcgagtt 900

DB 1016 cagccacagatcagaggaataatcgactgtcttcacgattgcccagtcgagtt 1075

QY 901 tacctcagtcctgacaccccaaaaaatagattggagaaaaattttaagagatatcc 960

DB 1076 tacctcagtcctgacaccccaaaaaatagattggagaaaaattttaagagatatcc 1135

QY 961 ttttgcctcagggatcagacacgtggaccatttgggttttatctgctggaggttcc 1020

DB 1136 ttttgcctcagggatcagacacgtggaccatttgggttttatctgctggaggttcc 1195

QY 1021 ggagtggtgggttttcttctgtgttctgtgttctcagtcacacaaatgaggtctggtt 1080

DB 1196 ggagtggtgggttttcttctgtgttctgtgttctcagtcacacaaatgaggtctggtt 1255

QY 1081 gatgaattatgatcaccctgaacaggccttcacggttgcccgagtcgacgacagct 1140

DB 1256 gatgaattatgatcaccctgaacaggccttcacggttgcccgagtcgacgacagct 1315

QY 1141 aaggcgcagccagctgtgtgagggctccctcgaagcctgcaagcctcctctgtgag 1200

DB 1316 aaggcgcagccagctgtgtgagggctccctcgaagcctgcaagcctcctctgtgag 1375

QY 1201 aggaatagagggaatgaattcttccaaaacaaactagaactgcaaaagcaccctgacaca 1260

AACT77404 standard; cDNA: 3727 BP.
AACT77404;
08-FEB-2001 (first entry)
Human ORFX ORF2959 polynucleotide sequence SEQ ID NO:5917.
Human: open reading frame; ORFX: detection; cytostatic; hepatotropic;
vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
hypotensive; dermatological; immunosuppressive; antiinflammatory;
antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
neurodegenerative disorder; osteoarthritis; graft vs host disease;
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
cholesterol ester storage; systemic lupus erythematosus; infection;
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
bone damage; cartilage damage; antiinflammatory disease; coagulation;
thrombosis; contraceptive; ss.
Homo sapiens.
WO200058473-A2.
05-OCT-2000.
31-MAR-2000; 2000WO-US08621.
31-MAR-1999; 99US-0127607.
02-APR-1999; 99US-0127636.
05-APR-1999; 99US-0127728.
30-MAR-2000; 2000US-0540763.
(CURA-) CURAGEN CORP.
Shinkets RA, Leach M;
WPI: 2000-602362/57.
P-PSDB; ABA43195.
Novel nucleic acids and peptides derived from open reading frame X,
useful for treating e.g. cancers, proliferative disorders,
neurodegenerative disorders and cardiovascular disease -
Claim 5; Page 5094-5096; 5507pp; English.
AACT7446 to AACT77606 encode the proteins given in ABA40237 to ABA43397,
which represent the human ORFX open reading frames 1 to 3161. The ORFX
sequences have activities such as: cytostatic; hepatotropic; vulnary;
antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
antidiabetic; hypotensive; dermatological; immunosuppressive;
antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
antithyroid; and antianaemic. The sequences can be used for determining
the presence of or predisposition to, or preventing or treating
pathological conditions associated with an ORFX-associated disorder. The
nucleic acids can be used to express ORFX proteins in gene therapy
vectors. The proteins and nucleic acids may be used to treat cancers,
proliferative disorders, neurodegenerative disorders, osteoarthritis,
graft vs host disease, cardiovascular disease, diabetes mellitus,
hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
bacterial or fungal infection, malaria, autoimmune disorders, asthma,
allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
nocturnal haemoglobinuria, antiinflammatory disease; to enhance
coagulation; to inhibit thrombosis; and as a contraceptive.
Sequence 3727 BP; 1031 A; 889 C; 931 G; 875 T; 1 other;

Query Match 82.7%; Score 2940.2; DB 21; Length 3727;
Best Local Similarity 99.3%; pred. No. 0;
Matches 3005; Conservative 0; Mismatches 13; Indels 7; Gaps 5;
Qy 532 tgcggcgcgtgacggtggtgcgcacaaagagctccgcggcctgacgagagtgcatc 591
Db 4 tgcggcgcgtgacggtggtgcgcacaaagagctccgcggcctgacgagagtgcatc 63
Qy 592 gagaagtccaatcacgtcagcggcagccgggggtcccgagagcccccacacccgcc 651
Db 64 gagaagtccaatcacgtcagcggcagccgggggtcccgagagcccccacacccgcc 123
Qy 652 catcgccgcgcacagggagcagagcctgtgcgcagcccatcgcaagtctctctcc 711
Db 124 catcgccgcgcacagggagcagagcctgtgcgcagcccatcgcaagtctctctcc 183
Qy 712 cagccggcctgcgctcgtccttttaggaagagctgcagagatggggcctccgaagc 771
Db 184 cagccggcctgcgctcgtccttttaggaagagctgcagagatggggcctccgaagc 243
Qy 772 agcggctcttcagctccttcgagagagcgacattgagaccactcattagcggacac 831
Db 244 agcggctcttcagctccttcgagagagcgacattgagaccactcattagcggacac 303
Qy 832 aatattgtgcagccacagatcagaggaataatcgaaatctgtcttcacgattggccag 891
Db 304 aatattgtgcagccacagatcagaggaataatcgaaatctgtcttcacgattggccag 363
Qy 892 tctgaagttacctcagtcctgcacacacacacacacacacacacacacacacacacac 951
Db 364 tctgaagttacctcagtcctgcacacacacacacacacacacacacacacacacacac 423
Qy 952 gagatataccttttgcctcaggggcacacacgctggaccactttgggtttatctgtcgg 1011
Db 424 gagatataccttttgcctcaggggcacacacgctggaccactttgggtttatctgtcgg 483
Qy 1012 gagtcttcggagggtggcggctttcattttgtctgttacctgttcttcagtgcaaatgag 1071
Db 484 gagtcttcggagggtggcggctttcattttgtctgttacctgttcttcagtgcaaatgag 543
Qy 1072 gctctggtgatgaaattatgatgacctgaaacagcgttcacggtggcgcagtgcaag 1131
Db 544 gctctggtgatgaaattatgatgacctgaaacagcgttcacggtggcgcagtgcaag 603
Qy 1132 cagacagtaaggcgcagccagcgtgtgtgagggtgctgcccctgcaaaagcctgcacaa 1191
Db 604 cagacagtaaggcgcagccagcgtgtgtgagggtgctgcccctgcaaaagcctgcacaa 663
Qy 1192 ctctgtgagagatagaggggaatgaattcttccacacacacacacacacacacacacac 1251
Db 664 ctctgtgagagatagaggggaatgaattcttccacacacacacacacacacacacacac 723
Qy 1252 ctgacgacataaccaatcagagcagcagcactatttttgaagaggttcagaaaatgaga 1311
Db 724 ctgacgacataaccaatcagagcagcagcactatttttgaagaggttcagaaaatgaga 783
Qy 1312 ccgagaaatgagcagcagagagaaatgaattgattttcttcttctgagatgttatgaa 1371
Db 784 ccgagaaatgagcagcagagagaaatgaattgattttcttcttctgagatgttatgaa 843
Qy 1372 gagaacagaaagac 1431
Db 844 gagaacagaaagac 903
Qy 1432 gagaatattggaagtgaattaccacccagtcgactcattttaggttaggtatgtgtgaaa 1491
Db 904 gagaatattggaagtgaattaccacccagtcgactcattttaggttaggtatgtgtgaaa 963
Qy 1492 aacaagcaaaagagatcttttaacagagctctttagaagattttgtcccggggtaataa 1551
Db 964 aacaagcaaaagagatcttttaacagagctctttagaagattttgtcccggggtaataa 1023

Oy 1552 gccagagcctgcaggaaacactccatcagttggtgatctggtatgctcctctctagtaca 1611
|||||
Db 1024 gccagagcctgcaggaaacactccatcagttggtgatctggtatgctcctctctagtaca 1083
Oy 1612 ttaagtaacacagcaaaagagccatctgtgtgtgaaagaggagccttgccatctctgag 1671
|||||
Db 1084 ttaagtaacacagcaaaagagccatctgtgtgtgaaagaggagccttgccatctctgag 1143
Oy 1672 agtccctttaagctccctcggtcctctcgaggagacotgtctccagtgactcggaagatccatctc 1731
|||||
Db 1144 agtccctttaagctccctcggtcctctcgaggagacotgtctccagtgactcggaagatccatctc 1203
Oy 1732 ccagaagagccagctccgctgtgcccacagcagcgccttccaggagcgagcaaacacccctg 1791
|||||
Db 1204 ccagaagagccagctccgctgtgcccacagcagcgccttccaggagcgagcaaacacccctg 1263
Oy 1792 agtcaacttcccatacgatgccaggaacacctccacaacctgcccgggggtcccgggggtt 1851
|||||
Db 1364 agtcaacttcccatacgatgccaggaacacctccacaacctgcccgggggtcccgggggtt 1323
Oy 1852 tcgcaagggaaacttatgaggtatcaactcagtgagcacagagacgcctcatgaaacgaag 1911
|||||
Db 1324 tcgcaagggaaacttatgaggtatcaactcagtgagcacagagacgcctcatgaaacgaag 1383
Oy 1912 gactttgaatccaaagcaaacctctgtgtgattctgtgtgagctcctgtgaagaccctg 1971
|||||
Db 1384 gactttgaatccaaagcaaacctctgtgtgattctgtgtgagctcctgtgaagaccctg 1443
Oy 1972 aggcattctctgagcgagcagatattctcctcagtagtagccacccgcagagggcgtagat 2031
|||||
Db 1444 aggcattctctgagcgagcagatattctcctcagtagtagccacccgcagagggcgtagat 1503
Oy 2032 tcttccagcagatataagattatccagagctggagagcttcccacagatctcccttta 2091
|||||
Db 1504 tcttccagcagatataagattatccagagctggagagcttcccacagatctcccttta 1563
Oy 2092 gaaccagttgtgaagtggcctcttgcccacacagagcaaaagaaagacacatct 2151
|||||
Db 1564 gaaccagttgtgaagtggcctcttgcccacacagagcaaaagaaagacacatct 1622
Oy 2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgc-tgcttagaat 2210
|||||
Db 1623 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatactgc-tgcttagaat 1682
Oy 2211 ggagaagggaaatacagagctccaagcctctgaaatgatttgcagaaagcgcctgaa 2270
|||||
Db 1683 ggagaagggaaatacagagctccaagcctctgaaatgatttgcagaaagcgcctgaa 1742
Oy 2271 gctcgattatgaagaaattactcctctttaaagaagtaactacagtggtgggaaagat 2330
|||||
Db 1743 gctcgattatgaagaaattactcctctttaaagaagtaactacagtggtgggaaagat 1802
Oy 2331 gcttagcactccagaaagatcaaaaaataagtttgacatggaaaaaatgcactcggtgt 2390
|||||
Db 1803 gcttagcactccagaaagatcaaaaaataagtttgacatggaaaaaatgcactcggtgt 1862
Oy 2391 tggcgaaggtgtccacgtcatccacaggtgaaatctgaaatttctagctagacaatt 2450
|||||
Db 1863 tggcgaaggtgtccacgtcatccacaggtgaaatctgaaatttctagctagacaatt 1922
Oy 2451 ccactttaaacacagtttcccagcaaacagcagccaaagatgtgcatacaaaagact 2510
|||||
Db 1923 ccactttaaacacagtttcccagcaaacagcagccaaagatgtgcatacaaaagact 1982
Oy 2511 cttaagcagctgacttcccagcagcatggatcttattgaccttggcgaaacctttcc 2570
|||||
Db 1983 cttaagcagctgacttcccagcagcatggatcttattgaccttggcgaaacctttcc 2042
Oy 2571 tacacacccatacttctctgcccagcttggagcagggacagctatcgtttacaacatttt 2630
|||||
Db 2043 tacacacccatacttctctgcccagcttggagcagggacagctatcgtttacaacatttt 2102
Oy 2631 gaagccttacttctctagaccaggaagtggatattgccaaggtctcagctttagc 2690

Db 2103 gaagggcctactcactctctagaccaggaagtgggtatttgccaaaggtctcagctttgtgacg 2162
Oy 2691 aggcattttgtctcttccatagtgagggaagagcggtttaaaaatgctccaagtttctgag 2750
|||||
Db 2163 aggcattttgtctcttccatagtgagggaagagcggtttaaaaatgctccaagtttctgag 2222
Oy 2751 gtttgaacatggggtcgtcggaacacagtatcggccaagacatgattttttacagatccagat 2810
|||||
Db 2223 gtttgaacatggggtcgtcggaacacagtatcggccaagacatgattttttacagatccagat 2282
Oy 2811 gtaccagctctcgaggtgtgttctcatgattaccacagagacctctacaatcacctcgagga 2870
|||||
Db 2283 gtaccagctctcgaggtgtgttctcatgattaccacagagacctctacaatcacctcgagga 2342
Oy 2871 gacagagatcgggccccgcctctcagctgcccctgtgttctctaccacatggtttgctcaca 2930
|||||
Db 2343 gacagagatcgggccccgcctctcagctgcccctgtgttctctaccacatggtttgctcaca 2402
Oy 2931 gttcccgcctgggtatctgtagccagagcttcttgatattttttcagggaaacagaggt 2990
|||||
Db 2403 gttcccgcctgggtatctgtagccagagcttcttgatattttttcagggaaacagaggt 2462
Oy 2991 catatttaaagtggtctttaagctctgttgggaagccataaagcccttgattctgcagcatga 3050
|||||
Db 2463 catatttaaagtggtctttaagctctgttgggaagccataaagcccttgattctgcagcatga 2522
Oy 3051 aaacctagaacaccatgttgactttataaaaaagcacgtaccacccaccccttggcttgggtaca 3110
|||||
Db 2523 aaacctagaacaccatgttgactttataaaaaagcacgtaccacccaccccttggcttgggtaca 2582
Oy 3111 gatggaagaagaccatcaatcaggtatttgaattggaatggacatcgctaaacagtttacaagctta 3170
|||||
Db 2583 gatggaagaagaccatcaatcaggtatttgaattggaatggacatcgctaaacagtttacaagctta 2642
Oy 3171 tgaagttgagtlaccacgtctctcaagaagaacttatcgattctctctctcctcagtgacaa 3230
|||||
Db 2643 tgaagttgagtlaccacgtctctcaagaagaacttatcgattctctctcctcagtgacaa 2702
Oy 3231 ccaaaagatggataaaatagagaaaaaacacagcagcttaacgcaaaacaccccttgacct 3290
|||||
Db 2703 ccaaaagatggataaaatagagaaaaaacacagcagcttaacgcaaaacaccccttgacct 2762
Oy 3291 ccttgaaacagttgagctggcagaaatgtaggtccaaagccttgagccacattgagaa 3350
|||||
Db 2763 ccttgaaacagttgagctggcagaaatgtaggtccaaagccttgagccacattgagaa 2822
Oy 3351 gctcctgagcagtgagagcaagctgaagcagggccatgcttaccttagaactggagcggtc 3410
|||||
Db 2823 gctcctgagcagtgagagcaagctgaagcagggccatgcttaccttagaactggagcggtc 2882
Oy 3411 ggcctctgctcagacagctggagagagctgcggcgaggcgagcgagccccagcgccggga 3470
|||||
Db 2883 g--cctgctcagacagctggagagagctgcggcgaggcgagcgagccccagcgccggga 2940
Oy 3471 gcttgagtgacgcagcagcccgagccacggcgactgacagctctgcagggagagattgcaa 3530
|||||
Db 2941 gcttgagtgacgcagcagcccgagccacggcgactgacagctctgcagggagagattgcaa 2999
Oy 3531 caccatccccacactgtccaggcctt 3555
|||||
Db 3000 caccat--cacactgtccaggcctt 3022

RESULT 5
AAI59640
ID AAI59640 standard; cDNA; 3281 BP.
XX
AC AAI59640;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 3629.

QY 1782 aaacaccctgagtcacttcccacatcgaatgcaggagaaacctccacaaacctgcccgggggtc 1841
DB 1261 aaacaccctgagtcacttcccacatcgaatgcaggagaaacctccacaaacctgcccgggggtc 1320
QY 1842 cccgggggttcgaaagaaacttatgagtgatcaactcagtgagcacagagagcctcca 1901
DB 1321 cccgggggttcgaaagaaacttatgagtgatcaactcagtgagcacagagagcctcca 1380
QY 1902 tgaacgaaggaactttgaatccaaagcaaacatcttgggtgattctgtgggaactcctgt 1961
DB 1381 tgaacgaaggaactttgaatccaaagcaaacatcttgggtgattctgtgggaactcctgt 1440
QY 1962 gaagaccggaggatctcctgaggcagcagatattctccgagtagcaccccgagaa 2021
DB 1441 gaagaccggaggatctcctgaggcagcagatattctccgagtagcaccccgagaa 1500
QY 2022 ggcgtgcgattcttcacagcagatagaagattattcagagctggagagcttcccaccg 2081
DB 1501 ggcgtgcgattcttcacagcagatagaagattattcagagctggagagcttcccaccg 1560
QY 2082 atctccttagaacagttgtgaagatgggccccttggcccaccacagaggaagaa 2141
DB 1561 atctccttagaacagttgtgaagatgggccccttggcccaccacagaggaagaa 1620
QY 2142 aagacatctcgtgagctccagagctgtgcaaaagctattctcaacagatactgt 2201
DB 1621 aaggacatctcgtgagctccagagctgtgcaaaagctattctcaacagatactgt 1680
QY 2202 gcttagaatggagaagaaatcagaagctccaagctcctgaaatgatttgcggaacaa 2261
DB 1681 gcttagaatggagaagaaatcagaagctccaagctcctgaaatgatttgcggaacaa 1740
QY 2262 gcgcctgaagctgattatgaagaaattactcctctttaaagaagtaactacagtgtg 2321
DB 1741 gcgcctgaagctgattatgaagaaattactcctctttaaagaagtaactacagtgtg 1800
QY 2322 ggaagagctctgagcctcagagaagatcaaaatgaatttgaatggaagaaatgca 2381
DB 1801 ggaagagctctgagcctcagagaagatcaaaatgaatttgaatggaagaaatgca 1860
QY 2382 ctggctgttgggaaggtgtg-cacgctcatcccgaggtgaaatctggaaatttctag 2440
DB 1861 ctggctgttgggaaggtgtgcccacgtcatcccgaggtgaaatctggaaatttctag 1920
QY 2441 ctgagcaattccacttaaacacagtttccacgcaaacacagcagcccaagagtgccat 2500
DB 1921 ctgagcaattccacttaaacacagtttccacgcaaacacagcagcccaagagtgccat 1980
QY 2501 acaagaactctttaaagcagctgacttcccagcagcagctgattctattgaccttgggc 2560
DB 1981 acaagaactctttaaagcagctgacttcccagcagcagctgattctattgaccttgggc 2040
QY 2561 gaaccttctacacacccacttctctgcccagcttggagcaggagacagctatcgcttt 2620
DB 2041 gaaccttctacacacccacttctctgcccagcttggagcaggagacagctatcgcttt 2100
QY 2621 acaacatttgaagcctactacttcttagaccaggaagtgaggatattgccaaggtctca 2680
DB 2101 acaacatttgaagcctactacttcttagaccaggaagtgaggatattgccaaggtctca 2160
QY 2681 gcttctgtagcaggaatttgccttctcatatgagtgaggaagagcgttttaaaatgctca 2740
DB 2161 gcttctgtagcaggaatttgccttctcatatgagtgaggaagagcgttttaaaatgctca 2220
QY 2741 agtttctgattgttgacatggggtgcggaaacagatgctccagacatgattattttac 2800
DB 2221 agtttctgattgttgacatggggtgcggaaacagatgctccagacatgattattttac 2280
QY 2801 agatccagatgtaccagctctcaggttgccttcattgattaccacagagacaccttacaatc 2860
DB 2281 agatccagatgtaccagctctcaggttgccttcattgattaccacagagacaccttacaatc 2340

QY 2861 acctggagagagcagagatcgccccagcctctactaagctgccccctgggttctctcaacatgt 2920
DB 2341 acctggagagagcagagatcgccccagcctctactaagctgccccctgggttctctcaacatgt 2400
QY 2921 ttgctcaacagttcccgctgggattcgtagccagagttcttgatattgattttcttcagg 2980
DB 2401 ttgctcaacagttcccgctgggattcgtagccagagttcttgatattgattttcttcagg 2460
QY 2981 gaacagaggttcatttttaaaagtggttttaagttctgttgggaagccataaagccttgattc 3040
DB 2461 gaacagaggttcatttttaaaagtggttttaagttctgttgggaagccataaagccttgattc 2520
QY 3041 tgcagcatgaaacctagaacacatagttgactttataaaagcagcgtacccaaccttg 3100
DB 2521 tgcagcatgaaacctagaacacatagttgactttataaaagcagcgtacccaaccttg 2580
QY 3101 gcttgggtacagatgaaagaccatcaatcaggttatttgaatggacatcgctaaacagt 3160
DB 2581 gcttgggtacagatgaaagaccatcaatcaggttatttgaatggacatcgctaaacagt 2640
QY 3161 tacaagcttatgaagttgagttaccacgtccttcaagaagaacttactcgtattctctctc 3220
DB 2641 tacaagcttatgaagttgagttaccacgtccttcaagaagaacttactcgtattctctctc 2700
QY 3221 tcaagtgaacacaaagaatgataaaattagagaaacacacagcagcttagcgaacacaga 3280
DB 2701 tcaagtgaacacaaagaatgataaaattagagaaacacacagcagcttagcgaacacaga 2760
QY 3281 acctgacctcttgaaacagttgcaggtggcaaatggttagatccaaagccttggaggcca 3340
DB 2761 acctgacctcttgaaacagttgcaggtggcaaatggttagatccaaagccttggaggcca 2820
QY 3341 ccattgagaagctcctgagcagctgagagcaagctgaaagcaggccatgttaccttagaac 3400
DB 2821 ccattgagaagctcctgagcagctgagagcaagctgaaagcaggccatgttaccttagaac 2880
QY 3401 tggagcgttcggccctcgtcgcagacggtggagagctcggcgagcgagcagagccca 3460
DB 2881 tggagcgttcggccctcgtcgcagacggtggagagctcggcgagcgagcagagccca 2940
QY 3461 gcgaccgggagcctgagtgacgcagcccgcagcccgccacggcgactgacagctctgcagga 3520
DB 2941 gcgaccgggagcctgagtgacgcagcccgcagcccgccacggcgactgacagctctgcagga 3000
QY 3521 gagattgcaacacacatccacacatgctccaggcctt 3555
DB 3001 gagattgcaacacacatccacacatgctccaggcctt 3035

RESULT 6

AAV05886

ID AAV05886 standard; DNA; 4039 BP.

XX

AAV05886;

XX

01-JUN-1998 (first entry)

DT

Human Tbc-1 gene.

DE

XX

tre-2; BUB2; cdc16; mast cell; probe; hybridisation; subtraction method; mouse; transcription factor; differentiation; proliferation; human; ds;

KW

KW

XX

XX

OS

XX

Key

Location/Qualifiers

403..3828

CDS

/*tag=

a

/product= Tbc1_protein

XX

US5700927-A.

XX

23-DEC-1997.

PD

XX	23-DEC-1994;	94US-0363300.	
PF	XX		
XX	23-DEC-1994;	94US-0363300.	
PR	XX		
XX	XX	(CHIL-) CHILDRENS MEDICAL CENT.	
PA	XX		
PI	Richardson P, Zon L;		
XX	XX		
XX	WPI; 1998-062437/06.		
DR	XX	P-PSDB; AAW44777.	
XX	XX		
PT	DNA encoding Tbc1 polypeptide - useful for treating leukaemia		
XX	Claim 3; Fig 1A-B; 22pp; English.		
PS	XX		
CC	XX	This is the nucleotide sequence which encodes a Tbc1 (tre-2, BUB2 and cdc16) polypeptide, isolated from a bone marrow-derived mast cell DNA library. The screen was carried out using a probe generated by a subtraction method which compared mRNA expression in an undifferentiated mast cell line PB15 and cell line PG16 (PB15 cells transformed to express the murine GATA-1 transcription factor - a factor which controls the expression of genes involved in mast cell differentiation). Tbc1 encodes a protein involved in the coupling of cell proliferation to cell differentiation, which can be used to treat leukaemia (especially acute CC myelogenous leukaemia) by causing leukaemic cells to differentiate.	
XX	XX		
SQ	Sequence 4039 BP; 1042 A; 1096 C; 1060 G; 841 T; 0 other;		
Query Match 70.6%; Score 2510.4; DB 19; Length 4039;			
Best Local Similarity 83.7%; Pred. No. 0;			
Matches 2949; Conservative 0; Mismatches 536; Indels 37; Gaps			
Qy	1	atggaaccaataacattcacaggaagaaacatctgctt-cttaacagaggtctcggtgga	59
Db	292	atggaggcaatcacattcacagggaggagcatcgtttccctaacgaagctccgcgga	351
Qy	60	ttttggcctcagctggt--gggtccctcgctgtgattccctgcaccacatgcccatg	117
Db	352	cttcggttcagctggttggtgtcctctaccgctgcatctctcaccaactatgcccatg	411
Qy	118	ctgcctcgggtgttggtgagtgaggcgggaagactcagcaggcaggtccaccagaagaacctg	177
Db	412	ctgcgctgggtgttagccgaggtacgaagactcagtgccagtgctctaaaaggagcct	471
Qy	178	gtaaccaagaactcgcgcttgctgtttcacacctctggacctgagatgtgaacctgagcca	237
Db	472	agacaaaagcaagtcgcctcttgggtttccgctccgacgcggtgcgagcgcgacctg	531
Qy	238	ggagaagatcaacagtgggatccctcgatctattccagcatctttgagtgcaagcctcag	297
Db	532	gagaaaagtcaacatgggaccgcgtcactctgtccagcatctttgagtgcaagcctcag	591
Qy	298	cgtgttcacaaactgattcacacagtcgatgaccacaaagttaacttgcctgtctgattaag	357
Db	592	cgtgtgcacaaactgattcacacagtcacgacccgagctactttgctgtctcattaaag	651
Qy	358	gaagacgctgtcacccgcgagatctctgtatgtttcaagaagcgatgatcaacaaaa	417
Db	652	gaggatgctgcacagggcagagcctctgctatgttttaagcagacgacacacaaaa	711
Qy	418	gtgcctgagatcatcagctccatccgtcagcggggaagatcgcccgagagagagctg	477
Db	712	gtgcctgagatcatcagctccatccgtcagcggggaagattgcccgcaagaagagctg	771
Qy	478	caetgcccgtccgagttcgcagacacggttttccagaagaagttcgaggtgctctctcgccg	537
Db	772	cgttgccctccgagttcgacgatacctctgcacaaaagttcgaggtgctctctctgtgccc	831
Qy	538	cgcgtgacggtggcgcacaaaggctccgcgcgcgcctgtgacgcagagtgatcgagaag	597
Db	832	cgggtgactgtgctcacaaaggcccccacccgcactgattgacgagtgatcgagaag	891

Qy	598	ttaataacagctcagcgccagccg	ggggttcgagagagccccgcgccaacccgcccacatgcc	657
Db	892	ttaaccatgtgagctgtgctgcagaa	cggactgggaagcgcaccccggcagccatca	951
Qy	658	gcgccacagggagccagagcctg	cgccagggcccatcgcaagctctctcccgagcc	717
Db	952	gc-----gctggccccaggcccatg	cgcgaatctctctcagcct	993
Qy	718	ggcctgcgctcgtggcctttaga	agagctgcaggga tgggggcccctcogaagcagcgc	777
Db	994	ggactgcgctcgtggccttcagg	agaggttccagga cgtcctccgagtagcctccgagtagcacc	1053
Qy	778	ttcttcagctccttcgagagagc	gcacattgagaacccactcattagcgcgacacaaatt	837
Db	1054	tt-----tagctcctttgacaatg	acatagagaacccctcatcgg tggggcacaaatgtg	1107
Qy	838	gtgcagcccacagatatcgagg	aaaaatcgaaactatgctcttcacgattggccagctctgaa	897
Db	1108	gttcagcccacagacatggagg	gagaccgaactatgctgttcacgattggcccatctgaa	1167
Qy	898	gttaactcatcagctcctgcac	ccccaaaaatagcatatggagaaaaattttaaggagata	957
Db	1168	gtttactctatcagctcgtcac	ccccaaaagattgcactggagaaaaattttaaggagata	1227
Qy	958	tcctttgctctcagggcacatc	agacagctggaccactttgggtttatctgtctcgggagctct	1017
Db	1228	tccttttgctctcagggca tca	gacatgggaccactttgggtttatctgtctcggaagctgc	1287
Qy	1018	tc-----cggaggtggcgctct	tcattttgtctgttaagtgctttcaagtgcacaaatgag	1071
Db	1288	tcgggtggcgcgagctggcg	gtcttcattttgtctgttaagtgcttcagtgacaaatgaa	1347
Qy	1072	gctctgtgtgatgaaattatga	gacccctgaaaacagccttcacagtggtggcgagctgcag	1133
Db	1348	gctctgtgttacgagata tca	tgacttcgaagcagcgtttcacggtgagctgcggtgcag	1407
Qy	1132	cagacagctaaggcgccagccc	agctgtgtgagggtgcgcccctgcgaagcctgcacaa	1191
Db	1408	cagacggtaaaggcaacagccc	agctctgtgagggtgcgcccctgcgaagcctgcacaa	1467
Qy	1192	ctctgtgagggatagaggaa	tgaattcttccaaaaaaaactagaactgcgaagcac	1251
Db	1468	ctctgcgaaggatagaggaa	tgaattctataaaccaaataggaaactccagaagcac	1527
Qy	1252	ctgcagacattaaccaatcag	gagcagcgactattttgaagaggttcagaaaaatgaga	1311
Db	1528	ttgaccacactgaccaatac	gaagcagggccacatatcgcgaaggttcagaaaaatgaga	1587
Qy	1312	ccgagaattgagcgcgagaga	aatgaattgtattttcttctgagatgtttatatgaa	1371
Db	1588	ccaagaaacgacgcagcga	gagagaatgaattaatattttcttctgaggtgcgttatgaa	1647
Qy	1372	gagaaacagaaaacacatcc	atatgtggggagatgaagcagacatcgcagatgcacga	1431
Db	1648	gagaagcaaaaaggcagcc	acactgcgggcgcaaacgacacacaggtggcagca	1707
Qy	1432	gagaataattggaaagtgaat	taccacccagtgccactcgatttaggtagatagatgcgaaa	1491
Db	1708	gagaataattgggaatgacct	gcccacccagtgctagcgggttcagggttagattgcgtgaag	1767
Qy	1492	aacaaagccaaagagatat	ttttaacagagctcttttagaaagtattttgtcccggggtaataaa	1551
Db	1768	aacgagcaaaagaggtcct	tcaacagagtcctcagagagcatctgtcccggggtaataaa	1827
Qy	1552	gccagaggtctgcagggaac	tccatccagtgtgattctggatagctccccgtgtctagtaca	1611
Db	1828	gccagaggtctgcagggaac	attccgcaggtgtgattctggacagctccacttctagtact	1887
Qy	1612	ttaagtacaccagcaagagc	catctgtgttgaaaaaggagccttgcccatctctgag	1671
Db	1888	ctagataacacgacgagag	agctgtccatgtgtgacaaagagccttcccgtctctgag	1947

PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isoqai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
PR Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs
XX
PS Claim 8; SEQ ID 18465; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 2362 BP; 657 A; 581 C; 578 G; 546 T; 0 other;
XX
Query Match 59.4%; Score 2112; DB 22; Length 2362;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1444 agtgaattaccaccagtcgactcgatttagctagatgctgtaaaacaaagcaag 1503
DB 1 agtgaattaccaccagtcgactcgatttagctagatgctgtaaaacaaagcaag 60
QY 1504 agatctttaacagagctcttagaaagtatttgcctccgggttaataaaagcagcctg 1563
DB 61 agatctttaacagagctcttagaaagtatttgcctccgggttaataaaagcagcctg 120
QY 1564 caggaaacactccatcagtgatcgtgtagctcctgctcctgtagtacattagtaacacc 1623
DB 121 caggaaacactccatcagtgatcgtgtagctcctgctcctgtagtacattagtaacacc 180
QY 1624 agcaagagaccatctgtgtgtaaaaggagccttgcccatctctgagagctccttaag 1683
DB 181 agcaagagaccatctgtgtgtaaaaggagccttgcccatctctgagagctccttaag 240
QY 1684 ctctcggtcctcggagacactgtccagtactcgagagagatcattcaccagaagacca 1743
DB 241 ctctcggtcctcggagacactgtccagtactcgagagagatcattcaccagaagacca 300
QY 1744 gctccgctcgtcccgagcagccttcaggagcgagcaaacacccctgagtcacttcccc 1803
DB 301 gctccgctcgtcccgagcagccttcaggagcgagcaaacacccctgagtcacttcccc 360

QY 1804 atcgaatgcaggaacacctccacaacctgcccgggggtccccgggggttttcgcaagaaaa 1863
DB 361 atcgaatgcaggaacacctccacaacctgcccgggggtccccgggggttttcgcaagaaaa 420
QY 1864 ctatgaggtatcactcagtcagtcagcagagacgcctcatgaacgaaagactttgaatcc 1923
DB 421 ctatgaggtatcactcagtcagtcagcagagacgcctcatgaacgaaagactttgaatcc 480
QY 1924 aaagcaaacctcttggtgattctctggtgagtcctctgtgaagaccgagagcattcctgg 1983
DB 481 aaagcaaacctcttggtgattctctggtgagtcctctgtgaagaccgagagcattcctgg 540
QY 1984 aggcagcagatatctcctccgagtagccaccgccgagagcgctggtgattcttccacaga 2043
DB 541 aggcagcagatatctcctccgagtagccaccgccgagagcgctggtgattcttccacaga 600
QY 2044 tatgaagattattcagagctggagagcttccccacgactcctctttagaacagattgt 2103
DB 601 tatgaagattattcagagctggagagcttccccacgactcctctttagaacagattgt 660
QY 2104 gaagatgggcccctttggccccccaccagaggaagaaagagacatctgtagctccga 2163
DB 661 gaagatgggcccctttggccccccaccagaggaagaaagagacatctgtagctccga 720
QY 2164 gagctgtggcaaaaggctatttctcaacagatactgctgttagaatggagaggaat 2223
DB 721 gagctgtggcaaaaggctatttctcaacagatactgctgttagaatggagaggaat 780
QY 2224 cagaagctccagcctctgaaatgattgctgaacagcgcctgaagctgattatgaa 2283
DB 781 cagaagctccagcctctgaaatgattgctgaacagcgcctgaagctgattatgaa 840
QY 2284 gaattactcctgtcttaaaagaagtaactacagtggtggaaagagctgttagcactcca 2343
DB 841 gaattactcctgtcttaaaagaagtaactacagtggtggaaagagctgttagcactcca 900
QY 2344 ggaagatcaaaaataaagtgttgacatggaaaaatgactcgcctggttgggcaaggctg 2403
DB 901 ggaagatcaaaaataaagtgttgacatggaaaaatgactcgcctggttgggcaaggctg 960
QY 2404 ccactcctacccagagtgaaatctggaatttctagctgagcaattccaccttaaacac 2463
DB 961 ccactcctacccagagtgaaatctggaatttctagctgagcaattccaccttaaacac 1020
QY 2464 cagtttcccagcaacagcagcagcaagagtggtgccatacaaaagaactcttaagcagctg 2523
DB 1021 cagtttcccagcaacagcagcagcaagagtggtgccatacaaaagaactcttaagcagctg 1080
QY 2524 acttcccagcagcatgcgattcttattgaccttggcggaaccttctctacacaccatcac 2583
DB 1081 acttcccagcagcatgcgattcttattgaccttggcggaaccttctctacacaccatcac 1140
QY 2584 ttctctgccagcttggagcagcagcagctatcgtttacaacatttgaaggcctactca 2643
DB 1141 ttctctgccagcttggagcagcagcagctatcgtttacaacatttgaaggcctactca 1200
QY 2644 ctctagacacaggaagtgggatatgtgccaaagtctcagcttctgtagcagagcatttgcct 2703
DB 1201 ctctagacacaggaagtgggatatgtgccaaagtctcagcttctgtagcagagcatttgcct 1260
QY 2704 ctctcatagtgaggaagagcgttttaaaatgctcaagtttctctgattgttgcacatgggg 2763
DB 1261 ctctcatagtgaggaagagcgttttaaaatgctcaagtttctctgattgttgcacatgggg 1320
QY 2764 ctgcgaaacagatctcggccagacatgatttttacagatccagatgtaccagctctcg 2823
DB 1321 ctgcgaaacagatctcggccagacatgatttttacagatccagatgtaccagctctcg 1380
QY 2824 aggttgccttcattgattaccacagagacacctctacaatcaacctggagagcagagatcggc 2883
DB 1381 aggttgccttcattgattaccacagagacacctctacaatcaacctggagagcagagatcggc 1440
QY 2884 ccagcctctacgctgccccctggttctcaccatgtttgctcctcacagtttccgcgtggga 2943

|||||
Db 1441 cccagctctacgtgccccctgggttcctcaacatgtttgctcacagttccccgctggga 1500
Qy 2944 ttcgtagccagagctctttgatgatattttctcagggaacagaggttcataatttaaagt 3003
Db 1501 ttcgtagccagagctctttgatgatattttctcagggaacagaggttcataatttaaagt 1560
Qy 3004 gctttaagctctgttgggaagcacaataagcccttgattctgcagatgaaacctagaacc 3063
Db 1561 gctttaagctctgttgggaagcacaataagcccttgattctgcagatgaaacctagaacc 1620
Qy 3064 atagttgactttataaaagacagctaccacaccttgcttgggtacagatggaagaacc 3123
Db 1621 atagttgactttataaaagacagctaccacaccttgcttgggtacagatggaagaacc 1680
Qy 3124 atcaatcaggtatttgaatggaacatcgctaaacagttacaaagctttatgaagttagtac 3183
Db 1681 atcaatcaggtatttgaatggaacatcgctaaacagttacaaagctttatgaagttagtac 1740
Qy 3184 cagctcttcaagaagaacttatcgattctctctctctcagtcagtcagacaaagaatgat 3243
Db 1741 cagctcttcaagaagaacttatcgattctctctctctcagtcagtcagacaaagaatgat 1800
Qy 3244 aaattagaaaaaacacacagcagcttacgcaaaacagacaccttgacctctcttgaacagttg 3303
Db 1801 aaattagaaaaaacacacagcagcttacgcaaaacagacaccttgacctctcttgaacagttg 1860
Qy 3304 caggtgcaaatgtagatcaaaagccttgaggccaccattgagaagctccttgagcagt 3363
Db 1861 caggtgcaaatgtagatcaaaagccttgaggccaccattgagaagctccttgagcagt 1920
Qy 3364 gagagcaagctgaagcagggcctgcttaccttagaactgagcggctcgccctctcag 3423
Db 1921 gagagcaagctgaagcagggcctgcttaccttagaactgagcggctcgccctctcag 1980
Qy 3424 acgdtgagagctgcggcgagcgagcagagcccgagcagcgggagcctgagtcgacg 3483
Db 1981 acgdtgagagctgcggcgagcgagcagagcccgagcagcgggagcctgagtcgacg 2040
Qy 3484 cagcccgagccacagcggcgactgacagctctgagggagagattgcaacacccatccacac 3543
Db 2041 cagcccgagccacagcggcgactgacagctctgagggagagattgcaacacccatccacac 2100
Qy 3544 tgtccaggcctt 3555
Db 2101 tgtccaggcctt 2112

RESULT 8
AAS34787
ID AAS34787 standard; cDNA; 1422 BP.
XX
AC AAS34787;
XX
DT 04-DEC-2001 (first entry)
XX
DE cDNA encoding novel human neoplastic disease associated polypeptide #21.
XX
KW Human; neoplastic disease associated polypeptide; cancer; gene therapy;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasotropic; ss.
XX
OS Homo sapiens.
XX
PN WO20015163-A1.
XX
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01358.
XX
XX 31-JAN-2000; 2000US-0179065.
PR
04-FEB-2000; 2000US-0180628.
PR
24-FEB-2000; 2000US-0184664.
PR
02-MAR-2000; 2000US-0186350.
PR
16-MAR-2000; 2000US-0189874.
PR
17-MAR-2000; 2000US-0190076.
PR
18-APR-2000; 2000US-0198123.
PR
19-MAY-2000; 2000US-0205515.
PR
07-JUN-2000; 2000US-0209467.
PR
28-JUN-2000; 2000US-0214886.
PR
30-JUN-2000; 2000US-0215135.
PR
07-JUL-2000; 2000US-0216647.
PR
07-JUL-2000; 2000US-0216880.
PR
11-JUL-2000; 2000US-0217487.
PR
11-JUL-2000; 2000US-0217496.
PR
14-JUL-2000; 2000US-0218290.
PR
26-JUL-2000; 2000US-0220963.
PR
26-JUL-2000; 2000US-0220964.
PR
14-AUG-2000; 2000US-0224518.
PR
14-AUG-2000; 2000US-0224519.
PR
14-AUG-2000; 2000US-0225213.
PR
14-AUG-2000; 2000US-0225214.
PR
14-AUG-2000; 2000US-0225267.
PR
14-AUG-2000; 2000US-0225268.
PR
14-AUG-2000; 2000US-0225270.
PR
14-AUG-2000; 2000US-0225447.
PR
14-AUG-2000; 2000US-0225757.
PR
14-AUG-2000; 2000US-0225758.
PR
14-AUG-2000; 2000US-0225759.
PR
18-AUG-2000; 2000US-0226279.
PR
22-AUG-2000; 2000US-0226681.
PR
22-AUG-2000; 2000US-0226868.
PR
22-AUG-2000; 2000US-0227182.
PR
23-AUG-2000; 2000US-0227009.
PR
30-AUG-2000; 2000US-0228924.
PR
01-SEP-2000; 2000US-0229287.
PR
01-SEP-2000; 2000US-0229343.
PR
01-SEP-2000; 2000US-0229344.
PR
01-SEP-2000; 2000US-0229345.
PR
05-SEP-2000; 2000US-0229509.
PR
05-SEP-2000; 2000US-0229513.
PR
06-SEP-2000; 2000US-0230437.
PR
08-SEP-2000; 2000US-0231242.
PR
08-SEP-2000; 2000US-0231243.
PR
08-SEP-2000; 2000US-0231244.
PR
08-SEP-2000; 2000US-0231413.
PR
08-SEP-2000; 2000US-0231414.
PR
08-SEP-2000; 2000US-0232080.
PR
12-SEP-2000; 2000US-0232081.
PR
14-SEP-2000; 2000US-0232397.
PR
14-SEP-2000; 2000US-0232398.
PR
14-SEP-2000; 2000US-0232399.
PR
14-SEP-2000; 2000US-0232400.
PR
14-SEP-2000; 2000US-0232401.
PR
14-SEP-2000; 2000US-0233063.
PR
14-SEP-2000; 2000US-0233064.
PR
14-SEP-2000; 2000US-0233065.
PR
21-SEP-2000; 2000US-0234223.
PR
21-SEP-2000; 2000US-0234274.
PR
25-SEP-2000; 2000US-0234997.
PR
25-SEP-2000; 2000US-0234998.
PR
25-SEP-2000; 2000US-0235484.
PR
27-SEP-2000; 2000US-0235834.
PR
27-SEP-2000; 2000US-0235836.
PR
29-SEP-2000; 2000US-0236327.
PR
29-SEP-2000; 2000US-0236367.
PR
29-SEP-2000; 2000US-0236368.
PR
29-SEP-2000; 2000US-0236369.
PR
29-SEP-2000; 2000US-0236370.
PR
02-OCT-2000; 2000US-0236802.
PR
02-OCT-2000; 2000US-0237037.

QY	3204	tatcagattcctctcctctcagtgacacacaaagaatgataataattagagaaacacacag	3263
Db	799	tatcagattcctctcctctcagtgacacacaaagaatggataaattagagaaacacacag	858
QY	3264	cagcttacgcaaacagaaacctgacctccttgaaacagttgcaggtggcaaatgttaggat	3323
Db	859	cagcttacgcaaacagaaacctgacctccttgaaacagttgcaggtggcaaatgttaggat	918
QY	3324	ccaaagccttgagccaccattgagaagctcctgagcagtgagagcaagctgaagcagcgc	3383
Db	919	ccaaagccttgagccaccattgagaagctcctgagcagtgagagcaagctgaagcagcgc	978
QY	3384	catgcttaccttagaactgagcggctcgccctgtgcagacggtgaggaggtgcggcg	3443
Db	979	catgcttaccttagaactgagcggctcgccctgtgcagacggtgaggaggtgcggcg	1038
QY	3444	gcggagcgcagacccacgacgcgggagcctgagtgacgcagcagcccgacccagcggcgga	3503
Db	1039	gcggagcgcagacccacgacgcgggagcctgagtgacgcagcagcccgacccagcggcgga	1098
QY	3504	ctgacagctctgcaggagagattgcaacaccatcccacactgtccagcctt	3555
Db	1099	ctgacagctctgcaggagagattgcaacaccatcccacactgtccagcctt	1150
RESULT	9		
AAK94796			
ID	AAK94796	standard; cDNA; 2989 BP.	
XX	AAK94796;		
AC			
XX			
DT	06-NOV-2001	(first entry)	
XX			
DE	Human full-length cDNA, SEQ ID NO: 3914.		
XX			
KW	Human; full length cDNA; cDNA synthesis; oligo-capping; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1130094-A2.		
XX			
PD	05-SEP-2001.		
XX			
PF	07-JUL-2000; 2000EP-0114089.		
XX			
PR	08-JUL-1999; 99JP-0194486.		
PR	11-JAN-2000; 2000JP-0118774.		
PR	02-MAY-2000; 2000JP-0183765.		
XX			
PA	(HELI-) HELIX RES INST.		
XX			
PI	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;		
PI	Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;		
XX			
XX	WPI; 2001-524255/58.		
DR	P-PSDB; AAM93840.		
XX			
PT	830 Primers useful for synthesizing full length cDNA clones and their		
PT	use in genetic manipulation -		
XX			
PS	Claim 8; SEQ ID NO 3914; 1380pp + sequence listing; English.		
XX			
CC	The invention relates to primers for synthesising full length cDNA		
CC	clones. 830 cDNA molecules encoding a human protein have been		
CC	isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA		
CC	molecules have been determined. Primers for synthesising the full		
CC	cDNA are useful for clarifying the function of the protein encoded by		
CC	the cDNA. The full length clones were obtained by construction of full		
CC	length enriched cDNA libraries that were synthesised by the oligo-capping		
CC	method. The primers enable the production of the full length cDNA easily		
CC	without any special methods. The present sequence is a full length		
CC	human cDNA of the invention.		

PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
FA Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-465558/50.
DR P-PSDB; AAU21586.
DR
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid
PT arthritis
XX
XX Claim 4; SEQ ID No 29; 687pp; English.
PS
XX The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides (AAU21568-AAU21851), and cDNA and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AAG34767-AAS35050 represent cDNA sequences
CC encoding for the novel human neoplastic disease associated polypeptides
CC of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1269 BP; 362 A; 313 C; 277 G; 312 T; 5 other;

Query Match 26.58; Score 943.2; DB 22; Length 1269;
Best Local Similarity 98.78; Pred. No. 7.7e-247;
Matches 967; Conservative 3; Mismatches 8; Indels 2; Gaps 2;

Qy 1908 aaaggacttgaatccaaagcaaacattcttggtgattctgtgggactctgtgaagac 1967
|||
Db 292 aaggagacttgaatccaaagcaaacattcttggtgattctgtgggactctgtgaagac 351
|||
Qy 1968 ccgaggcattcttgaggagcagagatattctccagtagtaccaccgccagaaaggcgtg 2027
|||
Db 352 ccgaggcattcttgaggagcagagatattctccagtagtaccaccgccagaaaggcgtg 411
|||
Qy 2028 cgattctccagcagatgaagattattcagagctggagagcttccccacagatctcc 2087
|||
Db 412 cgattctccagcagatgaagattattcagagctggagagcttccccacagatctcc 471
|||

Qy 2088 tttagaacaggtttgtgaagatgggccccttttggccccccaccagagaaagagac 2147
|||
Db 472 tttagaacaggtttgtgaagatgggccccttttggccccccaccagagaaagagac 531
|||
Qy 2148 atctcgtgagctccgagagctgtgcaaaagcctattcttcaacagatactgctgcttag 2207
|||
Db 532 atctcgtgagctccgagagctgtgcaaaagcctattcttcaacagatactgctgcttag 591
|||
Qy 2208 aatggagaagaaaaatcagaagctccaaagcctctgaaatgatttgcgaacagcgctt 2267
|||
Db 592 aatggagaagaaaaatcagaagctccaaagcctctgaaatgatttgcgaacagcgctt 651
|||
Qy 2268 gaagctcgattatgaagaaattactcctgtcttaagaagaagtaactacagtggtggaaaaa 2327
|||
Db 652 gaagctcgattatgaagaaattactcctgtcttaagaagaagtaactacagtggtggaaaaa 711
|||
Qy 2328 gatgcttagcaactccaggaagatcaaaaattaaagtgttgacatggaaaaatgcactcggc 2387
|||
Db 712 gatgcttagcaactccaggaagatcaaaaattaaagtgttgacatggaaaaatgcactcggc 771
|||
Qy 2388 tgttgggcaagggtgtgcacgtcatcaccaggtgaaatctggaattcttagctgagca 2447
|||
Db 772 tgttgggcaagggtgtgcacgtcatcaccaggtgaaatctggaattcttagctgagca 831
|||
Qy 2448 attccaccttaaacaccagtttccacgaaacacagcagcaaaagatgtgccatacaaga 2507
|||
Db 832 attccaccttaaacaccagtttccacgaaacacagcagcaaaagatgtgccatacaaga 891
|||
Qy 2508 actctaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctt 2567
|||
Db 892 actctaaagcagctgacttcccagcagcatgcgattcttattgaccttgggcgaacctt 951
|||
Qy 2568 tcttacacaccatacttctgtcccaggttggagcgagcagcatctgctttacaacat 2627
|||
Db 952 tcttacacaccatacttctgtcccaggttggagcgagcagcatctgctttacaacat 1011
|||
Qy 2628 ttgaagcctactcactctctagaccaggaagtggatattgccaaggtctcagctttgt 2687
|||
Db 1012 ttgaagcctactcactctctagaccaggaagtggatattgccaaggtctcagctttgt 1071
|||
Qy 2688 agcaggcatttgccttctcatatgagtgagggaagcggtttaaagtctcagatttct 2747
|||
Db 1072 agcaggcatttgccttctcatatgagtgagggaagcggtttaaagtctcagatttct 1131
|||
Qy 2748 gatgttgacatggggtgcggaaaacagtatcggccagacatgatttttacagatcca 2807
|||
Db 1132 gatgttgacatggggtgcggaaaacagtatcggccagac-tgattattttacagatcca 1190
|||
Qy 2808 gatgtaccagctctcgaggttgccttcattgattaccacagagacacctcacatcacctgga 2867
|||
Db 1191 gatgtwccagctctcgargttgcttcatgattaccacagana-ctctaccatcacctgga 1249
|||
Qy 2868 ggagcagcagatcgcccca 2887
|||
Db 1250 agaacacgaaatcngccca 1269
|||
RESULT 11
AAH04450
ID AAH04450 standard; cDNA; 883 BP.
XX
XX AAH04450;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:1285.
DE
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
PN
XX

PD 07-FEB-2001.
XX 28-JUL-2000; 2000EP-0116126.
XX 29-JUL-1999; 99JP-0248036.
XX 27-AUG-1999; 99JP-0300253.
XX 11-JAN-2000; 2000JP-0118776.
XX 02-MAY-2000; 2000JP-0183767.
XX 09-JUN-2000; 2000JP-0241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 1; SEQ ID 1285; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 883 BP; 245 A; 221 C; 219 G; 193 T; 5 other;
SQ

Query Match 22.1%; Score 785; DB 22; Length 883;
Best Local Similarity 96.7%; Pred. No. 9.9e-204;
Matches 843; Conservative 0; Mismatches 23; Indels 6; Gaps 4;
QY 1444 agtgaattaccacccagtcgactctgatttagctagatgctgaaacaaagcaag 1503
Db 1 agtgaattaccacccagtcgactctgatttagctagatgctgaaacaaagcaag 60
QY 1504 agatctttaaagagcttttagaagaattttgtccgggggtaataaagccagagcctg 1563
Db 61 agatctttaaagagcttttagaagaattttgtccgggggtaataaagccagagcctg 120
QY 1564 caggaaacacccatcagtggtggtcgtggtatgctcctgctagctacattaaagtaacac 1623
Db 121 caggaaacacccatcagtggtggtcgtggtatgctcctgctagctacattaaagtaacac 180
QY 1624 agcaagagacatctgtgtgtgaaaggagggccttgcccatctctgagagctctttaa 1683
Db 181 agcaagagacatctgtgtgtgaaaggagggccttgcccatctctgagagctctttaa 240
QY 1684 ctctcggctctcggagagacctgtccagtgactcggagagagctatctccagagagcca 1743
Db 241 ctctcggctctcggagagacctgtccagtgactcggagagagctatctccagagagcca 300

QY 1744 gctcgcgtgtcgcgccagcaggccttccagagagcgagcaaacacccctgagtcacttcccc 1803
Db 301 gctcgcgtgtcgcgccagcaggccttccagagagcgagcaaacacccctgagtcacttcccc 360
QY 1804 atcgaatgcaggaacctccaaacctgcgcgggggtcccccgggggttttcgcaaaagaaa 1863
Db 361 atcgaatgcaggaacctccaaacctgcgcgggggtcccccgggggttttcgcaaaagaaa 420
QY 1864 ctatgaggtatcactcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1923
Db 421 ctatgaggtatcactcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 480
QY 1924 aaagcaaacacacatttgggtgattctgtgtgggactcctgtgaagaccgcggaggaattcctg 1983
Db 481 aaagcaaacacacatttgggtgattctgtgtgggactcctgtgaagaccgcggaggaattcctg 540
QY 1984 aggcagcagatattctccagtagtagccaccccgagaaagcggtgcgattcttccacagaga 2043
Db 541 aggcagcagatattctccagtagtagccaccccgagaaagcggtgcgattcttccacagaga 600
QY 2044 tatgaagattattcagagctggagagcttccccccagctctcctttagaaccagttgt 2103
Db 601 tatgaagattattcagagctggagagcttccccccagctctcctttagaaccagttgt 659
QY 2104 gaagatgggcccctttggccccccacccagagagaaagaaagacatctcgtgagctccga 2163
Db 660 gaagatgggcccctttggccccccacccagagagaaagaaagacatctcgtgagctnca 719
QY 2164 gaggc-tgtgtgcaaaaggctatttcttcaacagatctcgtcttagaatggagagagaaa 2222
Db 720 agcttctgtgcaaaaggctatttcttcaacagatctcgtcttagaatggagagagaaa 779
QY 2223 tcagaagctccagcc---tcgaaatgatttgcgtgaaacagccg-lgaagctcgatt 2278
Db 780 tcagaagctccagcc---tcgaaatgatttgcgtgaaacagccg-lgaagctcgatt 839
QY 2279 atgaagaataattcctcctctttaaagaagta 2310
Db 840 atgaagaataattcctcctctttaaagaagta 871
RESULT 12
AAK92006
ID AAK92006 standard; cDNA; 849 BP.
XX AAK92006;
XX AC AAK92006;
XX AC AAK92006;
XX DT 06-NOV-2001 (first entry)
XX DE Human cDNA 5'-end sequence, SEQ ID NO: 466.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-0114089.
XX PR 08-JUL-1999; 99JP-0194486.
XX PR 11-JAN-2000; 2000JP-0118774.
XX PR 02-MAY-2000; 2000JP-0183765.
XX PA (HELI-) HELIX RES INST.
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
XX

PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -

XX Claim 2; SEQ ID NO 466; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is the nucleotide
 CC sequence of the 5'-end of a cDNA provided in the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX SQ Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;

Query Match 16.2%; Score 574.2; DB 22; Length 849;
 Best Local Similarity 98.2%; Pred. No. 3.5e-146;
 Matches 642; Conservative 0; Mismatches 6; Indels 6; Gaps 6;

QY 2555 ttggcgaaacctttctacacacccatacttctgtccagcttggagcaggacagctat 2614

Db 198 tggggcgaaacctttctacacacccatacttctgtccagcttggagcaggacagctat 257

QY 2615 cgctttacaacatttgaagcgctactactcttctagaccaggaagtgggatattgccaag 2674

Db 258 cgctttacaacatttgaagcgctactactcttctagaccaggaagtgggatattgccaag 317

QY 2675 gtctcagcttggtagcaggaatttggcttcttctatagtgaggaagagcggtttaaaa 2734

Db 318 gtctcagcttggtagcaggaatttggcttcttctatagtgaggaagagcggtttaaaa 377

QY 2735 tgcacaagttctgatgtttgacatggggtcggaacacagatcgcgcagacatgatta 2794

Db 378 tgcacaagttctgatgtttgacatggggtcggaacacagatcgcgcagacatgatta 437

QY 2795 ttttacagatccagatgtaccagctctcgaggtgtgcttctatgattaccacagagacctct 2854

Db 438 ttttacagatccagatgtaccagctctcgaggtgtgcttctatgattaccacagagacctct 497

QY 2855 acaatcacctggagagcagagatcgcccgccagctctacgtcgcccttggttctca 2914

Db 498 acaatcacctggagagcagagatcgcccgccagctctacgtcgcccttggttctca 557

QY 2915 ccattgttgcctcagctcccgctgggattcgtagccagagcttcttgatatttttc 2974

Db 558 ccattgttgcctcagctcccgctgggattcgtagccagagcttcttgatatttttc 617

QY 2975 ttcaggaaacagaggtctatattttaaagtgcctttaaagctgtgttggaagccataagccct 3034

Db 618 ttcaggaaacagaggtctatattttaaagtgcctttaaagctgtgttggaagccataagccct 677

QY 3035 tgattctgcagcatgaaacctagaacacatagttgattcttataaaagacagctaccca 3094

Db 678 tgattctgcagcatgaaacctagaacacatagttgattcttataaaagacagctaccca 736

QY 3095 accttgcttggtacagatgaaagaccatcaatcaggta-tttgaaatggacatc-gc 3152

Db 737 accttgcttggtacagatgaaagaccatcaatcaggta-tttgaaatggacatc-gc 795

QY 3153 taaacagttacaagcttatg-aagttgagtacca-cgtcttcaagaagaacct 3204

Db 796 taaacagttacaagcttatg-aagttgagttaccacccgtctttaaagaagaacct 849

RESULT 13

AAK93666

ID AAK93666 standard; cDNA; 849 BP.

XX AAK93666;

XX 06-NOV-2001 (first entry)

XX Human cDNA clone representative sequence, SEQ ID NO: 2126.

XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

XX Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99JP-0194486.

XX 11-JAN-2000; 2000JP-0118774.

XX 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI; 2001-524255/58.

XX 830 Primers useful for synthesizing full length cDNA clones and their

XX use in genetic manipulation -

XX Example 11; SEQ ID NO 2126; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence was used as the
 CC representative sequence from a human clone which was used in
 CC homology searches to identify the clone.

XX Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX SQ Sequence 849 BP; 224 A; 195 C; 196 G; 231 T; 3 other;

Query Match 16.2%; Score 574.2; DB 22; Length 849;
 Best Local Similarity 98.2%; Pred. No. 3.5e-146;
 Matches 642; Conservative 0; Mismatches 6; Indels 6; Gaps 6;

QY 2555 ttggcgaaacctttctacacacccatacttctgtccagcttggagcaggacagctat 2614

Db 198 tggggcgaaacctttctacacacccatacttctgtccagcttggagcaggacagctat 257

QY 2615 cgctttacaacatttgaagcgctactactcttctagaccaggaagtgggatattgccaag 2674

Db 258 cgctttacaacatttgaagcgctactactcttctagaccaggaagtgggatattgccaag 317

QY 2675 gtctcagcttggtagcaggaatttggcttcttctatagtgaggaagagcggtttaaaa 2734

Db 318 gtctcagcttggtagcaggaatttggcttcttctatagtgaggaagagcggtttaaaa 377

QY 2735 tgcacaagttctgatgtttgacatggggtcggaacacagatcgcgcagacatgatta 2794

Db 378 tgcacaagttctgatgtttgacatggggtcggaacacagatcgcgcagacatgatta 437

QY 2795 ttttacagatccagatgtaccagctctcgaggtgtgcttctatgattaccacagagacctct 2854

Db 438 ttttacagatccagatgtaccagctctcgaggtgtgcttctatgattaccacagagacctct 497

Qy	2855	acaataacctggaggagcagacagagatcaggccccagcctctacgtgccccctgggtcctca	2914
Db	498	acaataacctggaggagcagacagagatcaggccccagcctctacgtgccccctgggtcctca	557
Qy	2915	ccatgtttgctcacagttcccgctgggattcgtagccagagctcttggatgatattttc	2974
Db	558	ccatgtttgctcacagttcccgctgggattcgtagccagagctcttggatgatattttc	617
Qy	2975	ttcagggaacagaggttcatttaaagtggtcttaagtctgttgggaagccataaagccct	3034
Db	618	ttcagggaacagaggttcatttaaagtggtcttaagtctgttgggaagccataaagccct	677
Qy	3035	tgattctgcagcatgaaacctgaaaccatagttgaactttataaaagcacgtaccaca	3094
Db	678	tgattctgcagcatgaaacctgaaaccatagttgaactttataaaagcacgtaccaca	736
Qy	3095	accttggcttggtacagatggaaaagaccatcaatcaggta-tttgaaatggacatc-gc	3152
Db	737	accttggcttgg-acagatggaaaagaccatcaatcangtatatttgnaatggacatcgg	795
Qy	3153	taaacagttcaagcttatg-adgttgagtacca-cgtccttcaagaagaactt	3204
Db	796	taaacagttcaagcttatgaaagttgagtnccaccgtccttttaagaagaactt	849
RESULT 14			
AAZ50904			
ID	AAZ50904 standard; DNA; 17590 BP.		
AC	AAZ50904;		
XX			
DT	31-MAY-2000 (first entry)		
XX			
DE	Human TBC-1 partial genomic DNA comprising 5' end sequence.		
XX			
KW	TBC-1; human; biallelic marker; chromosome 4; cell cycle regulator; SNP;		
KW	Single nucleotide polymorphism; tissue differentiation; prostate cancer;		
KW	linkage analysis; genetic map; detection; diagnosis; genotyping;		
KW	transgenic animal; screening; ds.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	misc_signal	1..2000	
FT		/*tag= a	
FT		/*note= '5', Regulatory region"	
FT	exon	2001..2077	
FT		/*tag= b	
FT		/*number= 1	
FT	intron	2078..12739	
FT		/*tag= c	
FT		/*number= 1	
FT	misc_feature	9391..9845	
FT		/*tag= d	
FT		/*note= "Amplicon 99-430"	
FT	primer_bind	9391..9408	
FT		/*tag= e	
FT		/*bound_moiety= "Primer B1"	
FT		/*note= "Amplification of amplicon 99-430"	
FT	primer_bind	9475..9493	
FT		/*tag= f	
FT		/*bound_moiety= "Primer D1"	
FT		/*note= "Microsequencing of marker 99-430-352"	
FT	misc_binding	9482..9506	
FT		/*tag= g	
FT		/*bound_moiety= "Probe P1"	
FT		/*note= "Detection of Biallelic marker 99-430-352"	
FT	misc_feature	9494	
FT		/*tag= h	
FT		/*note= "Ambiguity base 'R' corresponds to 'A' in allele-	
FT		and 'G' in allele-2 of biallelic marker 99-430-352"	
FT	primer_bind	complement (9495)..9513)	

```

FT      /*tag= i
FT      /bound_moiety= "Primer E1"
FT      /note= "Microsequencing of marker 99-430-352"
FT      complement (9828..9845)
FT      /*tag= j
FT      /bound_moiety= "Primer C1"
FT      /note= "Amplification of amplicon 99-430"
FT      12392..12373
FT      /*tag= k
FT      /number= 1 bis
FT      12374..12739
FT      /*tag= l
FT      /number= 1 bis
FT      12740..13249
FT      /*tag= m
FT      /number= 2
FT      13250..17590
FT      /*tag= n
FT      /number= 2
FT      XX
FN      WO200008209-A2.
XX      17-FEB-2000.
XX      06-AUG-1999; 99WO-IB01444.
XX      07-AUG-1998; 98US-0095653.
XX      (GEST ) GENSET.
XX      Blumenfeld M, Bougueleret L, Chumakov I;
XX      WPI; 2000-205736/18.
XX      New isolated human TBC-1 nucleic acids, useful for developing products
XX      for the diagnosis and treatment of disorders involving cell
XX      proliferation, particularly prostate cancer -
XX      Claim 1; Page 93-100; 166pp; English.
XX      The present sequence is the partial genomic DNA of human TBC-1 gene,
XX      comprising the 5' regulatory region, exons 1, 1bis and 2. TBC-1 gene is
XX      mapped to a candidate region of prostate cancer on chromosome 4. Single
XX      nucleotide polymorphism (SNP) is located within the biallelic marker
XX      region 99-430-352, localised in intron 1 of TBC-1 genomic DNA.
XX      TBC-1 gene is involved in the regulation of cell cycle and tissue
XX      differentiation in mammals. An alteration of TBC-1 sequence may be
XX      associated with a pathological condition, resulting in abnormal cell
XX      proliferation leading to cancer, e.g. prostate cancer. The biallelic
XX      markers can be used for generation of genetic maps, linkage analysis and
XX      association studies. TBC-1 sequence can be used for detection,
XX      diagnosis, genotyping, production of transgenic animals and screening
XX      of compounds for use in therapy.
XX      Sequence 17590 BP: 4760 A; 3776 C; 4104 G; 4919 T; 31 other;

```

	Query Match	11.9%	Score 422.6;	DB 21;	Length 17590;
	Best Local Similarity	95.8%;	Pred. No. 5.4e-104;		
	Matches 434;	Conservative 0;	Mismatches 19;	Indels 0;	Gaps 0
Qy	1	atgaaccaaataattcacaggaaggaacatctgcttctcaacagaggtctcggtggat	60		
Db	12833	atggaaccaaataattcacaggaaggaacatctgcttctcaacagaggtctcggtggat	12892		
Qy	61	tttggcctgcagctggtgggtccctgcctgctgattccctgaccacacatgccatgctcg	120		
Db	12893	tttggcctgcagctggtgggtccctgcctgctgattccctgaccacacatgccatgctcg	12952		
Qy	121	ccctggggttgctggtgcagagatcagcaggcagtcaccagaaaggaacctgta	180		
Db	12953	ccctgggttgctggtgcagatcgaagatcagcaggcaatccaccagaaaggaacctgta	13012		

QY 181 accaagcaagtcggcttgcgtttccacctctggactgagatgtaacctgagccagg 240
 Db 13013 accaagcaagtcggcttgcgtttccacctctggactgagatgtaacctgagccagg 13072
 QY 241 agaagtaacagtgggatccctcgatctattccagcatctttgagtgcaagcctcagcgt 300
 Db 13073 agaagtaacagtgggatccctcgatctattccagcatctttgagtgcaagcctcagcgt 13132
 QY 301 gttcacaaactgattcaacacagtcacgaccacgaagttacttctgtctgataaggaa 360
 Db 13133 gttcacaaactgattcaacacagtcacgaccacgaagttacttctgtctgataaggaa 13192
 QY 361 gacgctgtccaccggcagagatctctgtatgtgttcaaaaggcagatgatacaaaaagt 420
 Db 13193 gacgctgtccaccggcagagatctctgtatgtgttcaaaaggcagatgatacaaaaagt 13252
 QY 421 cctgagatcatcagctccatcctgcagcggcg 453
 Db 13253 agtgagatggagatcccaaaagactaaggtgtg 13285

RESULT 15
 AAZ15229
 ID AAZ15229 standard; cDNA; 696 BP.
 XX
 AC AAZ15229;
 XX
 DT 12-OCT-1999 (first entry)
 XX
 DE Human gene expression product cDNA sequence SEQ ID NO:3698.
 XX
 KW Human; gene; gene expression product; diagnosis; therapy; probe;
 KW detection; mapping; tissue typing; profiling; forensic; cancer;
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9338972-A2.
 XX
 PD 05-AUG-1999.
 XX
 PF 28-JAN-1999; 99WO-US01619.
 XX
 PR 03-APR-1998; 98US-0030666.
 XX
 PR 28-JAN-1998; 98US-0072910.
 XX
 PR 24-FEB-1998; 98US-0075954.
 XX
 PR 31-MAR-1998; 98US-0080114.
 XX
 PR 03-APR-1998; 98US-0080515.
 XX
 PA (CHIR) CHIRON CORP.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;
 PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
 PI Stache-Crain B, Sudduth-Klinger J, Williams LF;
 XX
 DR WPI; 1999-494092/41.
 XX
 XX Novel human genes and their expression products which are
 PT differentially expressed in different cell types
 XX
 PS Claim 1; Page 1311; 2479pp; English.

CC be used for a variety of purposes, e.g. detection of expression levels,
 CC mapping, tissue typing or profiling, forensics, genetic analysis and
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
 CC can be used for raising antibodies for experimental, diagnostic and
 CC therapeutic purposes. The polynucleotides may also be used to construct
 CC arrays for diagnostics (which may be used to determine function of an
 CC encoded protein); and to detect differences in expression levels between
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
 CC identify a genetic predisposition or susceptibility to a disease such as
 CC cancer). The polynucleotides of the invention are especially used in the
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
 CC and lung cancer. The polynucleotides can also be used to screen for
 CC peptide analogues and antagonists.

XX Sequence 696 BP; 189 A; 157 C; 178 G; 142 T; 30 other;

Query Match 10.2%; Score 362; DB 20; Length 696;
 Best Local Similarity 100.0%; Pred. No. 2.8e-88;
 Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3194 aagaagaacttatcgtattctctctctcctcagtgacacacaaagaatgataaataagaga 3253
 Db 59 aagaagaacttatcgtattctctctcctcagtgacacacaaagaatgataaataagaga 118
 QY 3254 aaaccaacagcagcttacgcaaaacagacacttgacctcttgacacagcttgagtgca 3313
 Db 119 aaaccaacagcagcttacgcaaaacagacacttgacctcttgacacagcttgagtgca 178
 QY 3314 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 3373
 Db 179 atggtaggatccaaagccttgaggccaccattgagaagctcctgagcagtgagagcaagc 238
 QY 3374 tgaagcaggccatgcttaccttagaactgagcgggtcgccctgctgacacagctgaggg 3433
 Db 239 tgaagcaggccatgcttaccttagaactgagcgggtcgccctgctgacacagctgaggg 298
 QY 3434 agctgcggcggcggagcagagacccagcagcgggagcctgagtgacacagcccgagc 3493
 Db 299 agctgcggcggcggcggagcagagacccagcagcgggagcctgagtgacacagcccgagc 358
 QY 3494 ccacggcggactgacacagctctgcaggagagattgcaacacacccacacactgtccaggcc 3553
 Db 359 ccacggcggactgacacagctctgcaggagagattgcaacacacccacacactgtccaggcc 418
 QY 3554 tt 3555
 Db 419 tt 420

Search completed: September 9, 2002, 19:18:41
 Job time: 17499 sec

CC The present invention describes a library of human polynucleotides
 CC comprising the sequences given in AAZ12532 to AAZ1779. Also described is
 CC a method of detecting differentially expressed genes correlated with the
 CC cancerous state of a mammalian cell, comprising detecting at least one
 CC differentially expressed gene product in a test sample from a cell
 CC suspected of being cancerous, where the gene product is encoded by one
 CC of the 5248 polynucleotide sequences given in AAZ12532 to AAZ1779. The
 CC polynucleotides can be used as a source of primers and probes, which can

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2002, 14:24:47 ; Search time 8477.27 Seconds
(without alignments)
8775.683 Million cell updates/sec

Title: US-09-762-311-3_COPY_171_3725
Perfect score: 3555
Sequence: 1 atggaaccaataacattcac.....tccacactgtccaggcctt 3555

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_to.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description

1	2751	77.4	3023	9	BC014529
2	2510.4	70.6	4039	6	I86429
3	2510.4	70.6	4039	10	MMU33005
4	2340	65.8	2576	9	AB029031
5	2112	59.4	2362	9	AB027355
6	1359.2	38.2	1995	4	BTA17923
7	1179.6	33.2	3051	10	BC004675
8	998.4	28.1	1788	9	AK057182
9	682.2	19.2	5922	9	AB011175
10	465.8	13.1	189811	30	AC009595
11	465.8	13.1	195108	9	AC021106
12	465.8	13.1	208318	2	AC108933
13	424.2	11.9	110000	2	AL390202_04
14	304.6	8.6	3214	3	DMU50542
15	304.6	8.6	4746	3	DME17919
16	284.4	8.0	189811	30	AC009595
17	284.4	8.0	208318	2	AC108933
18	252.4	7.1	163928	2	AC021335
19	252.4	7.1	180918	9	AC098680
20	245.6	6.9	17019	2	AC019886
21	245.6	6.9	161117	3	AC008311
22	245.6	6.9	168448	3	AC009346
23	245.6	6.9	297832	3	AE003602
24	180	5.1	163928	2	AC021335
25	141.8	4.0	3761	9	BC014111
26	141	4.0	221717	2	AC074357
27	139.2	3.9	158310	9	AL162571
28	139.2	3.9	183577	2	AC011164
29	132.8	3.7	3691	6	AX078376
30	132.8	3.7	4560	9	HSA011679
31	123.8	3.5	154	6	AX261661
32	119.4	3.4	14119	2	AC013898
33	119.4	3.4	164413	3	AC023718
34	119.4	3.4	303648	3	AE003486
35	114.6	3.2	2013	9	AF042345
36	114.6	3.2	7383	9	AF008915
37	101.8	2.9	30027	3	CELT24D11
38	98	2.8	20553	3	CBRG26M15
39	95.6	2.7	73803	8	NCB11B22
40	90	2.5	5870	10	MMU53586
41	84.2	2.4	441	11	G47994
42	80	2.3	27102	3	U41540
43	79.6	2.2	110032	2	AC014916
44	79.6	2.2	159660	3	AC010040
45	79.6	2.2	282267	3	AE003555

ALIGNMENTS

RESULT 1
BC014529
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

BC014529
Homo sapiens
BC014529
Homo sapiens, clone IMAGE:3843156, mRNA.
BC014529.1 GI:17939551
human.

ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3023)
Strausberg,R.
Direct Submission
Submitted (24-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 14 Row: C Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
FEATURES
source Location/Qualifiers
1.3023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3843156"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH MGC-21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
BASE COUNT 883 a 712 c 740 g 688 t
ORIGIN

Query Match
Best Local Similarity 77.4% Score 2751; DB 9; Length 3023;
Matches 2751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 805 attgagaccacccattagcgacacataattgtcagcccccacagatcgcgagaaat 864
Db 10 ATTCAGAACCCCTATTAGCGGACACATATTGTGACGCCACAGATATCGAGGAAAT 69
QY 865 cgaactatgctcttcacgattggccagctggaagtttaacctatcagctcctgacaccaa 924
Db 70 CGAATATGCTCTTCAGGATGGCCAGCTGAAGTTTACCTCATCAGTCTCTGACACCA 129
QY 925 aaatagcatggagaaatttaagagagatatccttttctcagggagtgccgagcagc 984
Db 130 AAATATAGCATTTGGAGAAATTTTAAAGAGATATCTTTTGTCTCTCAGGCGATCAGAC 189
QY 985 gtgaccacttgggtttatctgtcgggagctcttcggaggtcctccgaggtggcgcttctt 1044
Db 190 GTGACCATTTTGGGTTATCTCTCGGGAGCTTCTCGGAGGTGGCGCTTTCATTGTGTC 249
QY 1045 tgtacgtgttcagtcgacaaatgggctctggttgatgaattatgatgacctgaaa 1104
Db 250 TGTACGTGTTTCAGTGCACAAATGAGGCTCTGTTGATGAATTTATGATGACCTGAAA 309
QY 1105 caggcttcacggtggccgagtcagcagacagcagcagcagcagcagcagcagcagcagc 1164
Db 310 CAGGCTTTCAGGTGGCGGAGTGCAGCAGACAGCTAAGGCGCCAGCCAGCTGTGTGAG 369
QY 1165 ggtgccccctgcaagcctgcacagctctgtgagagatagaggaatctctcc 1224
Db 370 GGCTGCCCTGCACAAACCTGCAACAGCTGTGTGAGAGGTAGAGGAATGAATCTTCC 429
QY 1225 aaacaaactagaaactgcaaaagccctgacgacattaccaaatacagcagcagcagc 1284
Db 430 AAACAAACTAGAACTGCACAAAGCACTGACGACATTAACCAATCAGGAGCAGCGACT 489
QY 1285 attttgagaggttcagaaattgagaccgagaaatgagcagcagcagcagcagcagcagc 1344
Db 490 ATTTTGAAGAGTTTCAAGAAATTTGAGCCGAGAAATGAGCAGCGAGAGAAATGAATTGAT 549
QY 1345 atttctttctgagatgtttatatgagagaaacagaaagacacatccattattggggag 1404
Db 550 ATTTCTTTTCTGAGATGTTTATATGAAGAAACAGAAAGAACACATCCATATTGGGGAG 609
QY 1405 atgaagcagacatcgagatggcagcagaggaatttggagtgaaattaccacccagtcg 1464
Db 610 ATGAGCAGACATCGCAGATGCGCAGAGAAATATTGGAAGTGAATTAACCCAGCTGCC 669

QY 1465 actcgatttagcgtagatatgctgaaaaacaaagcagaagagatctttaaagagagagag 1524
Db 670 ACTCGATTAGCTAGATATGCTGAAAAACAAGCAAGAGAGATCTTTAACAGAGTCTTTA 729
QY 1525 gaaagtatttttcccgggggttaaaagccagagggcctcaggaacacacccatcagtg 1584
Db 730 GAAAGTATTTTCCCGGGGTAAATAAGCCAGAGGCTTCAGAGACACATCCATCAGTGTG 789
QY 1585 gatctggatagcctcctgtctctagacatttaagtaaacacagcagaagagcctcgtgtg 1644
Db 790 GATCTGATAGTCCCTGTCTAGTACATTAACTAACACAGCAAGAGCCATCTGTGTG 849
QY 1645 gaaaagaggccttgcccatctctgagagctcctttaaagctcctcgtcctcgtcgtgag 1704
Db 850 GAAAAGGAGGCTTCGCCATCTCTGAGAGCTCTTTAAGCTCTCTCGGCTCTCTCGGAGG 1764
QY 1705 ctgtccagtgactcggagagtcctccagagagccagctcctcgtcgtcgtcgtcgtg 1824
Db 910 CTGTCCAGTGACTCGAGAGTCTATCTCCAGAGAGCCAGCTCCGCTGTCTCGCCAGCAG 969
QY 1765 gccttcaggaggcgaacaaacccctgagtcacttccctccatcgaatgccaggaacccca 1824
Db 970 GCTTTCAGGAGGCGAGCAACACCTGAGTCACTTCCCATCGAATGCCAGAACCTCCA 1029
QY 1825 caactcctccgggggtcccgggggttcgaaagggaacttatgaggtatcactcagtg 1884
Db 1030 CAACCTGCTCCGGGGTCCCGGGGTTTCGAAAGGAACTTATGAGGTATCACTCAGTG 1089
QY 1885 agcacagagacgctcatgaacgaagagcttgaatccaaagcgaacacacatctctgtg 1944
Db 1090 AGCAGAGAGCGCTCATGAACGAAGAGACTTTGAATCCAAAGCAAAACCATCTTGTGT 1149
QY 1945 tctgtgtgactcctgtgaagaccgggagcttctccagagatgaagatgggaccttggccc 2124
Db 1150 TCTGTGGGACTCTGTGAAGACCCGGAGGATTCCTCGAGGCGACAGATATTCCTCGA 1209
QY 2005 gtagcaccctccgagaaggcgtcgattcttcagagatgaagatgggaccttggccc 2184
Db 1210 GTAGCCACCCCGAGAGAGGCGTGGATTCCTCAGCAGATATGAAGATTTATTCAGAG 2064
QY 2065 gtagccttcccccagatcctctttagaaccagttgtgaaagatgggaccttggccc 2269
Db 1270 GGAGAGCTTCCCGCAGATCTCCTTTAGAACCAGTTTGTGAAGATGGGCCCTTGGCCC 2124
QY 2125 ccacagaggaaagaaaggacatctcgtgagctccgagagctgtggcaaaaggtctatt 2184
Db 1330 CCACAGAGGAAAGAAAGGACATCTCTGAGCTCCGAGAGCTGTGGCAAAAGGCTATT 1389
QY 2185 cttaacagatactgctgtcttagaatgagaggaataatcagaagctccagcctctgaa 2244
Db 1390 CTTCACAGATACTGTCTTAGAATGGAGAGGAAATCAGAAAGCTCCAGGCTCTGAA 1449
QY 2245 aatgattgtctgaacagcctgaagctcgtattgaagaaattactcctcgtctttaa 2304
Db 1450 AATGATTTGCTGAACAGCGCTGAAGCTGATTTAGAGAAATTTACTCCTCTCTTAA 1509
QY 2305 gaagtaactacagtggtgggaaagatgcttagcactccaggaagatcgaataattagtt 2364
Db 1510 GAAGTAACATACAGTGTGGGAAAGATGCTTAGCATCCAGGAGATCAAAAATTAAGTT 1569
QY 2365 gacatgggaaataatgactcgtcgtgttggaaggtgtgcacgtcatccagaggtgaa 2424
Db 1570 GACATGGAAATAATGCACCTCGCTGTTGGGCAAGGTGTGCACGTCATCACCAGGTGAA 1629
QY 2425 atctggaaattctagctgagcaattccacttaaacacaggtttcccgaggaacagcag 2484
Db 1630 ATCTGGAATTTCTAGCTGAGCAATTCACCTTAAACACAGTTTCCCGAGCAAAACAGC 1689
QY 2485 ccaagagatgctccatcaaaagaaactctaaagcagctgacttcccagcagatcgatt 2544
Db 1690 CCAAGGATGTCCCATACAAAGACTCTTAAAGCAGCTGACTTCCAGCAGCATGCGATT 1749

QY 778 ttcttcagctccttcgagagagcgacattgagaaacacacctcattagcggacacaaatt 837
Db 1054 TT-----TAGCTCTTGTACATGACATAGACAGAACCCACCTCATCGGTGGGCAATGTG 1107
QY 838 gtcagcccacagatcgcgagaaatcgaactatctcttcagatctggccagtcgtgaa 897
Db 1108 GTTCAGCCCCACAGACATGGAGGAGAACCGAACTATGCTGTTCAGGATTGGCCCATCTGAA 1167
QY 898 gtttacctatcagtcctgcacacaaataatagcattggagaaaaattttaaggagata 957
Db 1168 GTTTACTCTCATAGTCTCTGACACCAAAAGATTGCACTGGAGAAAAATTTAAGGAGATA 1227
QY 958 tccttttgcctcagggccatcagacacgctggaccactttgggtttatctgtcgggagtc 1017
Db 1228 TCCTTTTGTCTCTCAGGGCATACACATGTGGACCATTGTGGATTCTCGCCGAGAGTGC 1287
QY 1018 tc-----cggaggtggcgcttcttcattttgtctgttaagtggtttcagtcgacaaatgag 1071
Db 1288 TCGGGTGGCGGAGTGGCGCTTTTCATTTTGTCTGTACGTGTTCAGTGCACAAATGAA 1347
QY 1072 gctctggttgatgaattatgacccctgaacacagccttcacggtggccgagtcgag 1131
Db 1348 GCTCTGTTGACAGATCATGATGACTCTGAACAGCGCTTTCACGTAGCTGGGTGCGAG 1407
QY 1132 cagacagctaaggcgccagccagctgtgtgaggggtgccccctgcaaaagcctgcacaag 1191
Db 1408 CAGACGGCTAAGGCACGACGCGCTGTGTGAGGGCTGCCCTTTCGAAGCGCTGCACAAG 1467
QY 1192 cctctgagagatagaggaatgaattcttccaaacaaactagaactgcaaaagcac 1251
Db 1468 CTCTCGAAAGGATAGAGGAATGAATTCTATTAACCAAAATTAGAATCCCAAGAGCAC 1527
QY 1252 ctgacacattacacacacagcagcgagcagctatttttgaagaggtttcagaaattgaga 1311
Db 1528 TTGACCACACTGACCAATCAGGAGCAGCGCCACCATATTTCGAGGAGTTTCAGAAATTGAGA 1587
QY 1312 ccgagaaatgagcagcgagaaatgaattgattatttcttcttcagatgatttatgaa 1371
Db 1588 CCAGAAAGCAGCAGCAGAGAGAAATGAATTAATTTCTTTCTGTAGGTCTTATATGAA 1647
QY 1372 gagaaacagaaagacacacacattattgggagagatgaagcagacatcgagatggcaga 1431
Db 1648 GAGAACCAAAAGAGCAGCACAGCCACACTGTGGGGCGCCAAAGACACACTACAGGTGGCAGCA 1707
QY 1432 gagaatattggaagtgaattacacccagtcgcaactcgcaacttaggttagatgattatgtaa 1491
Db 1708 GAGAAATATTGGGAGTGACCTGCGCACCCACTGCTAGCCGGTTTCAGGTTAGATTTCGCTGAG 1767
QY 1492 acaaaagcaagagatcctttaacagagctctttagaagatttttgcgggggtaataaa 1551
Db 1768 AACAGAGCAAAAGAGTTCCTTAACAGAGTCCCTTAGAGAGCATTTCTGTCCGGGGTAATAA 1827
QY 1552 gccagagcctcaggaacacacacacagtggtatctgtagatgtagtccctgtctagtaca 1611
Db 1828 GCAGAGGGCTCGAGACCATTCGCGCAGTGTGGATCTGGACAGTCCACTTCTTAGTACT 1887
QY 1612 ttaagttaacaccagcaagagcactctgtgtgtaaaagagggccttgcccatctctgag 1671
Db 1888 CTAAGTAACACCAGCAGAGCTGTCCATGGGTGACAGGAGGCGCTTCCCGCTCTCTGAG 1947
QY 1672 agctcctttaagctcctcctcagagacacacacacacacacacacacacacacacac 1731
Db 1948 ACCTCTTCAAGCTCTTGGCTTCTCAGATGACCTGTCCAGTGACTCAGAGGGCCACATT 2007
QY 1732 ccagaagagcagctccgctgtgcgcccagcagggccttcagagagcgagcaaaacccctg 1791
Db 2008 GCAGAAAGAGTCTGCGCTGTGTTCACCCAGCAGGCGGTTTCAGAAAGAGAGGCCAACCCCTG 2067
QY 1792 agtcacttccccatcgaatgagcagaaactccacacacacacacacacacacacacacac 1851
Db 2068 AGTCATTTCCAGAGTAGAGTGGCCCTCGGCCCTCCAGAACTTCGCCAGAGCTCTCCAGGGGTC 2127
QY 1852 tcgcaaaaggaaacttatgaggtatcactcagtgagcagagagcgcctcattgaaagaaag 1911

Db 2128 TCTCAAAAGGAACTCATGCGGTACCCTCGTGAGACACAGAGAGCGCTCATGAACGCAAG 2187
QY 1912 gactttgaattccaaagcaaaacacacacattcttggtgattcttggtgggactcctctggaagaccgg 1971
Db 2188 GACTTTGAATCCAAAGCAAAACCACTTGGGTGACACAGATGGGACCCCGGTGAAGACCCGG 2247
QY 1972 aggcattcctcggagcagcagagatattcctccagtagtagcaaccgccgagaaagcggtgcgat 2031
Db 2248 CGGCACTCGTGGGAGACAGCAGATATTCTTCGAGTGGCCACTCCACAGAAAGGCTTTGTGAC 2307
QY 2032 tctccagcagagatagaagattattcagagctgggagagcttcccccacagatctccttta 2091
Db 2308 TCCCCGAGCAGATATGAAGATTATTCCGAGCTGGGAGAGCTCCCTCCACGCTCCCTTTA 2367
QY 2092 gaaccagtttgtgaagatggggcccttggcccccccccagcagagagaaagaaagaaacatct 2151
Db 2368 GAACCGGTGTGAGACGCGGCCAATTTGG-CCAGTACAGGAAGAAAGAGAGAGCTCA 2426
QY 2152 cgtgagctcgcagagcgtgtggcaaaaggctattcttcaacagatctg-ctgcttagaat 2210
Db 2427 CGCGAGCTTCGAGAGCTGTGGAAAAGGCCATCTTCGACGAGATCTCTGCTCCTCAGGAT 2486
QY 2211 ggaaagggaaatcagaagctccaagcctctgaaatgatttgcgaaacagcgccgtgaa 2270
Db 2487 GGAGAAAGGAGAAATCAGAAGCTACAAGCCTCTGAAAACGATTTTGTGTAACAACACGCTCA 2546
QY 2271 gctcgattatgaagaaatcactccctgtcttaaaagaaagtaactacagtggtgggaaagat 2330
Db 2547 GCTTGACTATGAAGAAATCCTCCGTGTCTTAAGAAGTCACTACAGTGTGGGAAAGAT 2606
QY 2331 gcttagcactccaggaagatcaaaaataagtttgacatggaaaaaatacactcggctgt 2390
Db 2607 GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGACATGGAAAAGACTGCATCACTGT 2666
QY 2391 tgggcaaggtgtgccacgtcatccagcaggtgaaactctggaattctcagtcgagcaatt 2450
Db 2667 TGGGCAAGGTGTGCCACGTCATCACCGAGGTGAGATCTGGAATAATTTAGTGAGCAGTT 2726
QY 2451 ccaccttaaacaccagtttccccagcaaacagcagcaaaaggatgtgccatacaaaagact 2510
Db 2727 CCACCTTAACACCCCATTTCTTAGTAACAGCAGCCAAAGAGCGTGCCTCTACAAAGAGCT 2786
QY 2511 cttaagcagctgacttccacagcagcagctgattcttattgaccttggcgaaaccttcc 2570
Db 2787 CCTGAAGAGCTGACCTGCCAGCAGCAGCCCATTTCTATCGAGCTCGGGGGAACCTTCC 2846
QY 2571 tacacacccatacttctctgccagcttggagcagagacagctatcgctttacaacatttt 2630
Db 2847 AACACATCCATCTCTCTGCCAGCTTGGAGCAGGTCACTGCTCACTTTACAACATTTCT 2906
QY 2631 gaagcctactcactcttagaccaggaagtggaatatgtccaaaggtctcagctttgtagc 2690
Db 2907 GAAGCGCTACTCGCTTCTGGACCAGGAGGTGGATTAAGTCCCAAGGTCTCAGCTTTGTGC 2966
QY 2691 aggcatttgcctctcatatagtgaggaagagcgtttaaaatgctcaagtctctgat 2750
Db 2967 AGGCATTTGCTTCTTCATCAGTGTGAGGAAGAGCGGTTCAGATGCTCAAGTTCCTGAT 3026
QY 2751 gttgacatggggctgcgaaacagatcgcgacagacatgattattttacagatccagat 2810
Db 3027 GTTTCACATGGGGCTGCGGAAACAGATCGGGCCAGACATGATTATTTTGCAGATCCAGAT 3086
QY 2811 gtaccagctctcagaggttgcctcatgattacacagagacacctctcaactcactcagagga 2870
Db 3087 GTTACAGCTGTACGGCTCTCCACGATTAACACCGAGACCTCTCAACACCTGGAAGA 3146
QY 2871 gcaagagatcgccccagcctctcagctgccccctgggttcctccacatggttgcctcaca 2930
Db 3147 GCACAGAGACTGGCCCCCTTACGTACGCGGCTCCCTGGTTTCTCACCGGTGTTCGCCCTCACA 3206
QY 2931 gttccgctgggattcgtagccagagctcttgatgatattttcttcaggggaacagaggt 2990

```
Db 3207 GTTCCCACTCGGCTTTGTAGCCAGAGCTTTTGATATGATCTTCTCCTTCAGGGATCAGAGGT 3266
Qy 2991 catatttaaagtgggttaagctctgttgggaagccataagcccttgattctgcagcatga 3050
Db 3267 CATATTTAAAGTAGCTTTAAAGTCTTTGGGGAGCCATAAGCCCTTGATTCACAGCATGA 3326
Qy 3051 aaactagaacaacatagttgactttataaaaagcagctaccacaacctgcttggtaca 3110
Db 3327 GAACCTGGAAACCATCGTGGAGCTTCATGAAGAACACACTCCCAACCTGGGCTGTGCTGA 3386
Qy 3111 gatgaaaagaccatcaatcagtgatttgaatggacatcgctaaacaggttacaagctta 3170
Db 3387 GATGAGAAGACCATCAGTCAGTGTGTTGAGATGGACATCGCAAGCAGCTCCAGGCCCTA 3446
Qy 3171 tgaagttgagtcacagctctctcaagaagaacttatcagattctctctctcagtgacaa 3230
Db 3447 TGAGTTCGAGTACCAAGCTCGTCAGGAGGAGCTTATGAGTCTCGCCTCTCAGTGACAA 3506
Qy 3231 ccaagaatggataaattagagaacacacacagcagcttacgcaacacagaaaccttgact 3290
Db 3507 CCAGAATGGAGAATTTGGAGAAAACCCACAGCAGCTTGCCCAACAGAACCTTGACCT 3566
Qy 3291 cettgaacagttgtaggtggcaaatggttaggtccaaagccttgaggccaccattgagaa 3350
Db 3567 CCTGGAGCAGTTGAGGTGGCAATGCTAGGATCCAAAGCCTTGAAGCCAGCGTAGAGAA 3626
Qy 3351 gctctgagcagtagagcaagcagctgaagcagcagcagcttacttagaactgagcggctc 3410
Db 3627 ACTTCTTACCACGAGAGTAAGCTGAAGCAGCTGCGCTGACCTGGAGGTGGAGCG--T 3684
Qy 3411 ggcctgtgcagacggtgagagagctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 3470
Db 3685 CGCCTGTGTCAGATGCTGGAGAGAGCTGCGGAGGCAAGCGCCGCGCCAGCAGCTCCAGA 3744
Qy 3471 gctgtagcagcagccgagcccccagcggcgagctgacagct 3512
Db 3745 GCAGACTGCACCCAGCTGGAGCCACAGCAGCGATTGACCGCT 3786

RESULT 3
MMU33005 house mouse.
LOCUS Mus musculus tbc1 mRNA, complete cds. 4039 bp mRNA linear ROD 12-JAN-1996
DEFINITION U33005
ACCESSION U33005
VERSION U33005.1 GI:988220
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4039)
AUTHORS Richardson,P.M. and Zon,L.I.
TITLE Molecular cloning of a cDNA with a novel domain present in the
tre-2 oncogene and the yeast cell cycle regulators BUB2 and cdc16
JOURNAL Oncogene 11 (6), 1139-1148 (1995)
MEDLINE 96032578
PUBMED 7566974
REFERENCE 2 (bases 1 to 4039)
AUTHORS Richardson,P.M. and Zon,L.I.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-1995) Leonard I. Zon, Hematology, Childrens
Hospital of Boston, HHMT, 300 Longwood Ave., Boston, MA 02115, USA
FEATURES
source
location/Qualifiers
1..4039
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="mast cells"
1..4039
/gene="tbc1"
403..3828
/gene="tbc1"
/feature="ORF1"
/codon_start=1
```

```
/product="Tbc1"
/db_xref="GI:988221"
/translation="MPMLPVWVAEVRRLSGCCKEPTKQVRLWVSPSGURCBPDLLE
KSQWPLDICSIFECQPKQVHKLHNSHDPYSFACLIKEDAAHRSQCYVFKADQDT
KVEPIISSIRQAGKIAQOECPSEFDDTFAKKEFVLFCCGRVTVAHKAPALIDEC
LEKHNHVSGRTDWEAPTQGPSAPGPPMRKSFOPGLRSLAFKRFQDASLRSTF
SSFUNDIENHLIGGHNVPQPTDMEENTMLTIGPSEVYLSPDTKTALSKNKEIS
FCSGIRHVDHFGICRECSGGSGGHFVCIYVQCETNEALVDEIMTKLQATVAAV
QCTAKAPQCEGCPLOGLHKLCSRIEGBMNSKTKLEQKHLTLTTNOQATIFEVQ
LDRPNREORENELISFURLYEKQESHSTGAPKQLQVAANESDLPSPASRFR
LDSLKNRAKRLTESLESLSRGNKARGLOHDSVLDLSTSTLSNTSKLSMGDK
EAPVSTSFKLIGSDSDLSDEGTHAESLSPQAFRRRANTLSHFVPECAPP
EPASSQVGSOKLWVHVSSTPEHKKDEFESKANHLGDTGTPVTRHSWRQIIF
LRVATPOKACDSPSEYEDYSELGELPPRSPLEPCVEDGPGFQYKKRRRRASFPSCG
KPSCSRSCLVMEKENOKLASENLDLKLKLDYIEITFCLKEVITVWKMSTPG
RSKTKFDMEKVHSAVGQVPRHGEITWKLAEQFHLKPPFPKQKQDVKYKELUK
LTSQHAIDLGTFTPTHTYFSAQAGQLSLYNILKAYSLLDQEVGYCGLSVAG
ILLHMBEFAFKMLKFLDMGLRKQYRPMIILQIQMYQLSRLLHGYHRYLHLE
EHENGPTTAAFWFLTPASQPLGFVARVDMIFLOQSEVIFKVALSLLSHKLPLIL
PHENLETIVDFIKNTLNLGLVQMEKTIQVPEMDIAKQLOAYEVHVQEELESS
LDSNQRMEKLEKNTSLRKONLDLQLOVANARIQSLEATVEKLLTSEKLRQAL
TLEVRPPAADGGGAEEAKRPAQRHLPAGARRHLTAARRDCAPLTLSKP"
BASE COUNT 1042 a 1096 c 1060 g 841 t
ORIGIN
Query Match 70.6%; Score 2510.4; DB 10; Length 4039;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 2949; Conservative 0; Mismatches 536; Indels 37; Gaps 8;
Qy 1 atggaacaaataacattcacagcaaggaacacatcgctt-ctaacagaggtctcggtga 59
Db 292 ATGGAGGCANTACATTCACAGCAGGAAGAGCATCGTTCCCTACGAGAGTCTCCGTGGA 351
Qy 60 ttttgccctgcagctggt--gggctccctgctgtgcatctccctgaccaccatgccatg 117
Db 352 CTTTCGGCTTGCAGCTGGTGGTCTCCTTACCCGTCATCTCTCACCACACTTACCCCATG 411
Qy 118 ctgacctgggtgtggtgagtgagagacatcagcagcagcagtcacaccagaagaagacat 177
Db 412 CTGCGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 471
Qy 178 gtaaccaagaagtcgctgttgccttccacctctgagctgagatgagtgagcagcagcag 237
Db 472 AGGACAAAGCAAGTCCGCTTTGGGTTTCGCCCTCGGACTCGGCTGCGAGCTGACCTG 531
Qy 238 gggagaagtcaacagtgaggatccctctgattctccagcatctttgagtcgaagcctcag 297
Db 532 GAGAAAGTCAACCATGGAGCCCGCTCATCTCTTCCAGCATCTTTCAGTGCAGGCTCAG 591
Qy 298 cgtgttcaaaactgatttcaacacagtcacacacacacacacacacacacacacacacac 357
Db 592 CGTGTGCACAACTGATTACACACAGTACACGACCCGAGCTACTTTGCTTGCCTCATTAAG 651
Qy 358 gaagacgctgtccacgcagcagatgctgctgattgttcaagcgcgagtgatcaacacaaa 417
Db 652 GAGGATGTCGCACAGCGCAGAGCCTCTCTATGTTGTTTAAAGCAGACAGCATCAACAAA 711
Qy 418 gtgctgagatcatcagctccatccgctcagcggcgggaagatcgcccgagagagctg 477
Db 712 GTGCTGAGATCATCAGCTCCATCCGCGAGCGCGGAGAGATTGCCCGCCAGGAAGAGCTG 771
Qy 478 cactgcccgtccgagttcagacacagcttttccagaagattcgaggtgcttcttcgcgc 537
Db 772 CGTTGCCCTCCGAGTTCGACGATACCTTCGCCAAAAAGTTTCAGGTGCTCTTCTGTGTC 831
Qy 538 cgcgtgacggtggcgacaagaagctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 597
Db 832 CGGGTGACTGTGGCTCACAAGAAAGGCCCCCACCCTGACCTGTTGACAGAGTGTATCGAAG 891
Qy 598 ttcaatcacgtcagcggcagcggggggtccgagagagccccgcacacacccgcacccatgcc 657
```

Db 892 TTCACCATGTGAGCTGTGTCGCAGACGGACTGGAGCGCCACCGGGCAGCCATCA 951
Qy 658 gcccccaagggagccagggagcctgtgctgagcccatgcccgaagtccttctccagccc 717
Db 952 GC-----GCTGGCCCCAGGCCCATGGCAAAATCTCTTCACAGCCT 993
Qy 718 ggcctgcctgcctgccccttttagaaggagctcagagatgggggctccgaagcagggc 777
Db 994 GGAAGTGGCTGCTGGCTTTCAGAGAGGAGTCCAGAGCGCTAGCCCTCCGCGATAGCACC 1053
Qy 778 ttcttcagctcccttcgagggagcgacattgagaaacccatctattagcgacacaaatt 837
Db 1054 TT-----TAGCTCTTGTGACATGACATAGAGAACCCCTCATCGTGGGACATGTG 1107
Qy 838 gtccagcccaagatatcgaggaaaaatcgaaactatgctcttcagattggccagtcgtgaa 897
Db 1108 GTTCAGCCCAACAGACATGGAGGAGAACCGAACTATGCTGTTTCAGATTTGGCCCATCGAA 1167
Qy 898 gtttacctcagctcctcagacacaaataatagcattggagaaaaattttaaggagata 957
Db 1168 GTTTTACCTATCATGCTCTGACACCAAAAGATTGCACTGGAGAAAATTTTAAGGAGATA 1227
Qy 958 tctctttgctcctcagggcatcagacacgtggaccactttgggtttatctgctcgggagtct 1017
Db 1228 TCCTTTTGTCTCAGGGCATCAGACATGTGGACCCTTTGGATTTCATCTGCCGAGAGTGC 1287
Qy 1018 tc-----cggaagtgccgcttctctttgtctgttaagtgcttcagtcgacaaatgag 1071
Db 1288 TCGGGGTGGCGGAGTGGCGCTTTTCATTTTGTCTGTACGTGTTCAGTGCACAAATGAA 1347
Qy 1072 gctctggttgatgaattatgagccctgaaacagccttcacgctggccgagtcgag 1131
Db 1348 GCTCTGTTGACAGATCATGATGACTCTGAGCAGGCTTTACAGTAGTGGGTGCGAG 1407
Qy 1132 cagacagtaagggccagcccgctgctgtaggggtgccccctgcaaaagcctgcacaag 1191
Db 1408 CAGACGCTAAGGCACACGAGCCAGCTCTGTGAGGGGTGCCCTTGCAGAGCCTGCAACAG 1467
Qy 1192 cctgtgagaggatagaggaatgaattctccaaacaaactagaaactgcaaaagcac 1251
Db 1468 CTCTGCGAAAGGATAGAGGAATGAATTCATCTAAACCAAAATAGAACTCCAGAGCAC 1527
Qy 1252 ctgacgacattaaccatcaggcgaggcgactatttttgaagaggttcagaaattgaga 1311
Db 1528 TTGACCACACTGACCAATCAGAGCAGCGCCACCATATTCGAGGAGGTTTCAGAAATGAGA 1587
Qy 1312 ccgagaaatgagcagcagagaaatgaattgattattctttctgagatgtttatgaa 1371
Db 1588 CCAAGAAACAGCAGCGAGAGAAATGAATTAATTAATTTCTTTTCTGAGGTGCTTATATGAA 1647
Qy 1372 gagaacagaaagacacatccattatggggagatgaagcagacatcgagatggcagca 1431
Db 1648 GAGAGCAAAAGACAGCACAGCCACACTGGGGGCCCAAGCAGACACTACAGGTGGCAGCA 1707
Qy 1432 gagaatttggaagtgaattaccaccacagtgccactcgtatttaggtagatatgctgaa 1491
Db 1708 GAGAATATTGGGAGTGACCTGCCACCCAGTGCTAGCCGGTTACGTTAGATTGCTGTAAG 1767
Qy 1492 acaagcaaaagagatctttaacagagcttttagaagatattttgtcccggggtaataaa 1551
Db 1768 AACAGAGCAAAAGAGTCTTTAACAGAGTCCCTAGAGAGCATTTCTGCCGGGGTAATAA 1827
Qy 1552 gccagggcctgcagaaacactccatcagtggtgattctgagatgctccctgtctagtaca 1611
Db 1828 GCCAGAGGCTCGAGACCAATTCGCCAGTGTGGATCTGGACAGCTCCACTTCTAGTACT 1887
Qy 1612 ttaagtaacaccagcaaaagccatctgtgtgaaaggaggccttgccatctctgag 1671
Db 1888 CTAAGTAACACCAGCAAGAGCTGTCCATGGGTGACAAAGGAGGCTTTCCCGGCTCTGAG 1947
Qy 1672 agtctcttaagctcctcggtcctcctcgagagacactgtccagtgactcgagagctcctc 1731
Db 1948 ACCTCCTTAAGCTCTTGGCTCCTCAGATGACCTGTCCAGTGACTGTCCAGTGACATCCAGT 2007

Qy 1732 ccagaagagccagctccgctgtctgccccagcaggtcccttcagggagcgagcaaacacccctg 1791
Db 2008 GCAGAGAGTCTGCCCTGTGTGTACCCCCAGCAGCGCTTCAGAAAGAGAGCAACACCCCTG 2067
Qy 1792 agtcacttccccatgaattgcaggaacctccaacctgcccggggtccccgggggtt 1851
Db 2068 AGTCATTTCCAGTAGAGTGCCTGCGCTCCAGAACCTTCGCCAGAGCTCTCCAGGGGTC 2127
Qy 1852 tcgcaaggaacacttatagggtatcactcagtgagcacagagacgcctcatgaacaaaag 1911
Db 2128 TCTCAAAGGAAACTCATCGGTACCACTCCGTGAGCACAGAGACGCTCATGAACGAAG 2187
Qy 1912 gactttgaatccaaagcaaacccatcttgggtgattcttgggtgactcctgtgaaagaccgg 1971
Db 2188 GACTTTGAATCCAAACAAACCACTGGGTGACACAGATGGGACCCCGCTGAAGACCCCG 2247
Qy 1972 aggcactctcgagcgagcagatattcctccagtagtcacaccccgacagaagcgctgcg 2031
Db 2248 CGGCACCTCTGGAGACAGCAGATATTCCTTCGAGTGGCCACTCCACAGAGGCTTGTGAC 2307
Qy 2032 tcttcagcagatatgaattatttcagagctgggagagcttccccacagatctctcttta 2091
Db 2308 TCCCCAGCAGATATGAAGATTATTCGAGCTGGGAGAGCTCCCTCACGCTCCCTTTA 2367
Qy 2092 gaacagtttgtgaagatggcccttggccccccacagaggaagaaagacacatct 2151
Db 2368 GAACCGGTGTGAGACCGGCCCATTTGG--CCAGTACAGGAAGAAAGAGGAAGAGCTCA 2426
Qy 2152 cgtgagctccgagagctgtggcaaaaggctattcttcaacagatctg-ctgcttagaat 2210
Db 2427 CGCGAGCTTCGAGAGCTGTGGAAAAAGGCCATCTTCAGCAGATCTGCTCGCTCAGGAT 2486
Qy 2211 gagaaggaataatcagaagctccaagcctctgaaaaatgatttgcgtgaaagcgctgaa 2270
Db 2487 GGAGAAGGAGAACTCAGAAGCTACAAAGCTGTGAAAACGATTTTCTGTAACAAACGCTCAA 2546
Qy 2271 gctcgattatgaagaaattactcctgtcttaagagtaactacagtgctgggaaagat 2330
Db 2547 GCTTGACTATGAAGAANTCCTCCGTGTCTTAAAGAAAGTCACTACAGTGTGGGAAAGAT 2606
Qy 2331 gcttagcactccagggaagatcaaaaaataagtttgacatggaaaaatgcaactcggtgt 2390
Db 2607 GCTTAGCACTCCAGGAAGATCCAAAATTAAGTTTGATGGAATAAGTGCACTACAGTGT 2666
Qy 2391 tgggcaagtgctgcacgtcaccagagtgaaatctgaaatttctagctgagcaatt 2450
Db 2667 TGGGCAAGGTGTGCCACGTCATCACCAGGTGAGATCTGGAAATTTCTAGCTGAGCAGTT 2726
Qy 2451 ccaccttaaacacacaggtttcccgagcaaacagcagccaaaggatgccccatacaaaagact 2510
Db 2727 CCACCTTAAACACCCATTTCTTAGTAACAGCAGCAAGCAAGGACGTGCCCTACAAAGAGCT 2786
Qy 2511 cttaagcagctgactcccgagcagatcgatctcttattgacctggcggaacctttcc 2570
Db 2787 CCTGAAGAGAGTCACTCGCAGCAGCAGCCCATTCCTACGAGCTCGGGGAAACCTTTCC 2846
Qy 2571 tacacacccatctctctgcccagcttgagcagcagcagctatcgctttacaacatttt 2630
Db 2847 AACACATCATCTCTCTGCCCCAGCTTGAGCAGGTGAGTACTGCCAAGGTCTCAGCTTTGTGGC 2906
Qy 2631 gaaggcctactcactcttagaccaggaagtggaattatgccaaaggtctcagcttttagc 2690
Db 2907 GAAGGCTTACTCGTCTTGACCAGGAGGTGGATACTGCCAAGGTCTCAGCTTTGTGGC 2966
Qy 2691 aggcatttgcctctcatatgagtgaggaagggcggtttaaaaatgctcaagtttctgat 2750
Db 2967 AGGCATTTTGTCTTCTTACATGAGTGAAGAGGCGTTCAAGATGCTCAAGTTCCTGAT 3026
Qy 2751 gtttgcacatgggctgcgaaacagtatcgccagacatgattattttacagatccagat 2810
Db 3027 GTTTGACATGGGCTCGGGAAACAGTATCGCCAGACATGATTTATTTTTCAGATCCAGAT 3086

QY	2811	gtaccagctctcaggttgcttcaatgattaccacagagacctctacaatcaccttgaggga	2870
Db	3087	gtaccagctctcaggttgcttcaatgattaccacagagacctctacaatcaccttgaggga	3146
QY	2871	gcacgagatcgccccagcctctacgtcgccccctggttctcaccatgtttgcctcaca	2930
Db	3147	gcacgagatcgccccagcctctacgtcgccccctggttctcaccatgtttgcctcaca	3206
QY	2931	gttccgcctcgaggttcagccagagctcttctgatatgattttcttcagggaacagaggt	2990
Db	3207	gttccgcctcgaggttcagccagagctcttctgatatgattttcttcagggaacagaggt	3266
QY	2991	cataattaaagtggccttaagtctgttggaagccataagcccttgatctcgagcatga	3050
Db	3267	cataattaaagtggccttaagtctgttggaagccataagcccttgatctcgagcatga	3326
QY	3051	aaactagaacacatagtgcatttataaaagcagcctaccacaccttgcttggtaca	3110
Db	3327	aaactagaacacatagtgcatttataaaagcagcctaccacaccttgcttggtaca	3386
QY	3111	gatggaagaacacatcaggttattgaaatggacatcgctaaacaggttacaagctta	3170
Db	3387	gatggaagaacacatcaggttattgaaatggacatcgctaaacaggttacaagctta	3446
QY	3171	tgaagtgtgatacaactccttcaagaagaacttatgatctctctcctcagtgacaa	3230
Db	3447	tgaagtgtgatacaactccttcaagaagaacttatgatctctctcctcagtgacaa	3506
QY	3231	ccaaagaatgataaattagagaaaacacacagcagcttaacgaaacacaccttgacct	3290
Db	3507	ccaaagaatgataaattagagaaaacacacagcagcttaacgaaacacaccttgacct	3566
QY	3291	ccttgacagcttcaggtggcaaatgtaggattccaaagccttgaggccaccttgagaa	3350
Db	3567	ccttgacagcttcaggtggcaaatgtaggattccaaagccttgaggccaccttgagaa	3626
QY	3351	gctcctgagcagtgagagcaagctgaagcaggccatgcttacccttagaactggagcgctc	3410
Db	3627	gctcctgagcagtgagagcaagctgaagcaggccatgcttacccttagaactggagcgctc	3684
QY	3411	ggccctgctcagacggtgagagctcgggcgagcgagcgagagccagcagcagggga	3470
Db	3685	ggccctgctcagacggtgagagctcgggcgagcgagcgagagccagcagcagggga	3744
QY	3471	gccttgagtcgacgagccgagccacggcgagctgacgct	3512
Db	3745	gccttgagtcgacgagccgagccacggcgagctgacgct	3786
RESULT	4		
LOCUS	AB029031	2576 bp	linear
DEFINITION	Homo sapiens mRNA for KIAA1108 protein, partial cds.		
ACCESSION	AB029031		
VERSION	AB029031.1		
KEYWORDS	GI:5689552		
SOURCE	Homo sapiens brain cDNA to mRNA, clone_lib:pbluescriptII SK plus clone:hk03387.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (sites) Kikuno, R., Nagase, T., Ishikawa, K., Hirose, M., Miyajima, N., Tanaka, A., Kotani, H., Nomura, N. and Ohara, O.		
TITLE	Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro		
JOURNAL	DNA Res. 6 (3), 197-205 (1999)		
MEDLINE	99397452		
REFERENCE	2 (bases 1 to 2576) Ohara, O., Nagase, T. and Kikuno, R.		
AUTHORS	Direct Submission		
TITLE	Submitted (17-JUN-1999) Osamu Ohara, Kazusa DNA Research Institute,		

Laboratory of DNA Technology; Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)

FEATURES
 source
 1. .2576
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /clone="hh03387"
 /tissue_type="brain"
 /clone_lib="pBluescriptII SK plus"
 1. .2292
 /gene="KIAA1108"
 <1. .2292
 /gene="KIAA1108"
 /note="hh03387 cDNA clone for KIAA1108 has a 3306-bp
 insertion."
 /codon_start=1
 /product="KIAA1108 protein"
 /protein_id="BAA83060.1"
 /db_xref="GI:5689553"
 /translation="NSSKTKLEQLKHLTLTNOEQATIFEEOVKLRPRNEQRENELII
 SFURCLVEEKKQKHIGEMQKTSQMAAENIGSELPSPTFRFLDKNNKAKSLTES
 LESILSRGNKARGQLQHSISVDLDSLSLTSNTSKSPSVCEKEALPISSSPFKLLGS
 SELDSSDSHLEPEAPLSPQAFRRRRANTLSHFPIEQEPPQARGSPGVSQRIKL
 RYHSVETETPHERKDESKANHIGDGGTVPKTRHSRQOIFLRVATPOKACDSSRK
 YEDYSELGPPRSPLEKEDGDFGPPPEEKRTSRRELWOKALQOILLRMKX
 ENGLQASNDENKFLKLDYERITPCLKEVTVWVKMLSTPGRSIKFDMERKHSVA
 GQGVPRHHGEINWKLAEOPHLKQPSKQOPDVPYKELKOLTQSOHAILLDLGRT
 FPHFYPFSAQLGQSLNLYLKAYSLIDOEVCYCGLSFVAGILLHSHSEFAFKML
 KFMDFMGKQRTFPMIDILQIQYQLSRLLHDYHRDLNHLSEHIGSPSLYAPFNL
 TMFASQFPGFVARVDMLTLOQTEVIFKVALSLDGHKPLLILOHENLETIVDFIKST
 LNLKGVQKMETNQVFEMDIKQLOAYEVEYHVLQEELIDSSPLSDNORMDKLETTN
 SPRLKQNLDDLEOLVANGRIQSLQLEATIEKLSSSESKLKQAMLTLELRSALLQTVEE
 LRRRSREPSDREPECTQPEPTG"

BASE COUNT 746 a 615 c 628 g 587 t

ORIGIN

Query Match 65.8%; Score 2340; DB 9; Length 2576;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2340; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1216 aattcttccaaaacaaactagaaactgaaagcacctgcagacattacacaaatcaggag 1275
 |||||
 Db 1 AATTCTTCCAAAACAAACTAGAACTGCAAAAGCACCTGCAGCATTAAGCAATCAGGAG 60
 |||||

Qy 1276 caggcgactattttgaagaggttcagaattgagaccgagagaaatgacgcgagagaat 1335
 |||||

Db 61 CAGCGCACTATTTTGAAGAGGTTTCAGAAATTGAGACCGAGAAATGAGCGGAGAGAA 120
 |||||

Qy 1336 gaattgattattcttctcagatgtttatgaagagaaacagaagaacacatccat 1395
 |||||

Db 121 GAATTGATATTTCTTTCTGAGATGTTTATTAAGAGAAACAGAAACACATCCAT 180
 |||||

Qy 1396 attggggagatgaagcagacatcgagatggcagcagagaatatgtgaagtgaattacca 1455
 |||||

Db 181 ATTGGGGAGATGAAGCAGACATCGCAGATGCGCAGCAGAGAAATATTGGAAGTGAATTACCA 240
 |||||

Qy 1456 cccagtgccactcgatttaggctagatatgctgaaaaacaaagcaagagagatctttaaca 1515
 |||||

Db 241 CCCAGTGCCCACTGCATTATAGGCTAGATATGCTGAAAACAAAGCAAGAGAGATCTTTAACA 300
 |||||

Qy 1516 gagctcttagaaaagtattttgtcccggggttaataaagccagagagcctgcaggaacactcc 1575
 |||||

Db 301 GAGTCTTAGAAAGTATTTTGTCTCCGGGGTAAATAAGCCAGAGGCGCTGCAGGAACACTCC 360
 |||||

Qy 1576 atcagtgatcgatcgatagctccctgtctagtacattaaagtaacacacagcaagacca 1635
 |||||

Db 361 ATCAGTGTGATCTGGATAGTCTCCTGTCTAGTACATTAACTAAACACACAGCAAGACCA 420
 |||||

Qy 1636 tctgtgtgtaaaaggaggccttggccatctctctgagagctcctttaagctcctcggtcc 1695
 |||||

Db 421 TCTGTGTGTGAAAGAGAGGCTTTGCCATCTCTGAGAGCTCCTTTAAAGCTCCTCGGCTCC 480
Qy 1696 tcggaggacctgtccagtgactcgagagtcattctcccaagagagccagctccgctgtcg 1755
Db 481 TCGGAGGACCTGTCCAGTGACTCGGAGAGTCATCTCCCAAGAGAGCCAGCTCCGCTGTG 540
Qy 1756 cccagcagccttcagagagcgagcaaacacccctgagtcacttcccatcaatgagcag 1815
Db 541 CCCAGCAGCCTTCAGGAGGAGCAACACCCCTGAGTCATTTCCCNATCGAATGCCAG 600
Qy 1816 gaacctccacaacctccgggggtcccccgggggttgcgaagaaacttatgaggtat 1875
Db 601 GAACCTCCACAACCTCCCGGGGGTCCCGGGGGTTTCGAAAGAAACTTATGAGGTAT 660
Qy 1876 cactcaagtacagacagacgcctcatgaacgaaggaactttgaatccaaagcaaacat 1935
Db 661 CACTCAGTGAGCAGAGAGCGCTCATGAACGAAGAGACTTTGAATCCAAAGCAAAACCAT 720
Qy 1936 ctgggtgattctgggtgagcctctggaagacccggagggcattcctggaggcagcagata 1995
Db 721 CRTGGTGATCTGGTGGGACTCTGTGAAGACCCGGAGGCATTCCTGGAGGCGAGATA 780
Qy 1996 ttctccgagtgaccccccgcagagggcgtgcgattctccagcagatatgaagattat 2055
Db 781 TTCTCTCGAGTAGCCACCCCGCAGAGGCGTGCATTTCTCCAGCAGATATGAAGATTAT 840
Qy 2056 tcagagctggagagcttcccccagcagatctctttagaaccagtttgaaagatggccc 2115
Db 841 TAGAGCTGGGAGAGCTTCCCCACGATCTCTTTAGAACCAGTTTGTGAATGGGCCC 900
Qy 2116 ttggccccccaccagagaaagaaagacatctcgtgagctccgagagctgtggcaa 2175
Db 901 TTTGGCCCCCACCAGAGGAAAGAAAGAGACATCTCGTGAGCTCCGAGAGCTGTGGCAA 960
Qy 2176 agggctattcttcaacagatactgctttagaagtgagaaagaaataatcagaagctccaa 2235
Db 961 AAGGCTATTCTTCAACAGATACTGCTGTAGTAATGGAGAAGGAAATCAGAAGCTCCAA 1020
Qy 2236 gctctgaaatgattctgaacagccctgaagctcgattatgaagaaattactccc 2295
Db 1021 GCCTCTGAANAATGATTGCTGAACAAGCCCTGAAGCTCGATTATGAANAATTAATCTCC 1080
Qy 2296 tgccttaagaagtaactacagctgtgggaaagatgcttagcactccaggaagatacaaaa 2355
Db 1081 TGTCTTAAAGAACTAATACAGTGTGGGAAAGATGCTTAGCACATCCAGGAAGATCAAAA 1140
Qy 2356 attaaattgacatgaaataatgactcggtgttgaggcaaggtgtgccaagtcatcac 2415
Db 1141 ATTTAAGTTTGACATGGAAAAATGCACTCGGCTGTGGGCAAGGTGTGCCACGTATCAC 1200
Qy 2416 cgaggtgaaatctggaaattctagctgagcaattccacttaaacaccagtttcccagc 2475
Db 1201 CGAGGTGAATCTGGAAATTTCTAGCTGAGCAATTCACCTTTAAACACCAGTTTCCAGC 1260
Qy 2476 aaacagcagccaaagatgtgcatacaagaactttaaagcagctgacttcccagcag 2535
Db 1261 AAACACAGCCAAAGGATGTGCATACAAAGAACTCTTAAAGCAGCTGACTTCCCAGCAG 1320
Qy 2536 catgcatcttattgacattggggaacatttctctacacaccataacttctctgcccag 2595
Db 1321 CATGCGATTTCTTATGACCTTGGGCAAGCTTTCTCTACACACCCATPACTTCTGTGCCAG 1380
Qy 2596 cttggagcagcagcagctatcgctttacaacatttgaagccctactcactcttagaccag 2655
Db 1381 CTTGGAGAGGAGAGCTATGCGCTTTTACAATTTTGAAGCCCTTACTTCTTAGACCAG 1440
Qy 2656 gaagtggatttgcgaagctcagctttgtagcagggcattttgtcttctcatatgagt 2715
Db 1441 GAAGTGGGATATGCCAAGGCTCAGCTTGTAGCAGGCAATTTGCTTCTCATATGACT 1500
Qy 2716 gaggaaagcgttttaaatgctcaagtttctgattgtgacatggggtgcggaaacag 2775
Db 1501 GAGGAAGAGCGGTTTAAATGCTTCAAGTTTCTGATGTTTACATGTTTACATGGGGTGGGGAACAG 1560

Qy 2776 tatcgccagacagctattatttaccagatccagatgtaccagctctcaggttgcattcat 2835
Db 1561 TATCGCCAGACATGATTATTTACAGATCCAGATGTACCAGCTCTCGAGGTGTCTTAT 1620
Qy 2836 gattaccacagagacctctacaatcacctcgaggagcagagatcgccccagcctctac 2895
Db 1621 GATTACCACAGAGACCTCTACAACTCACCTGGAGGACGAGATCGGCCCCAGCTCTAC 1680
Qy 2896 gctgccccctggttccctcacatgtttgcctcacagttccccctgggattctagccaga 2955
Db 1681 GCTGCCCTCGTGTCTCCTCACCATGTTGCCCTCACAGTTCCCGCTGGGATTCGTAGCCAGA 1740
Qy 2956 gctcttgatagattttcttcaggaacagagtcataatttaaagtggtcttaagtctg 3015
Db 1741 GTCCTTGATATGATTTTCTTCAGGAACAGAGGTCATATTTAAAGTGGCTTTAAGTCTG 1800
Qy 3016 ttgggaagccataaagcccttgattctgcagcatgaaacctagaaccatagttgacttt 3075
Db 1801 TTGGGAAGCATAAGCCCTTGATTCTGCAGCATGAAACCTAGAAACCATAGTTGACTTT 1860
Qy 3076 ataaaaagcagcctaccaccccttggttggtacagatggaaagacccatcaatcagta 3135
Db 1861 ATAAAAAGCACGCTACCCAACTTGCTGTGACAGATGGAAGAACCATCAATCAGGTA 1920
Qy 3136 ttgaaatggacatcgctaaacagttacaagcttatgaagttgagtcaccacgtccctcaa 3195
Db 1921 TTTGAANTGGACATCGCTAAACAGTTACAAGCTTATGAAGTTGAGTACCACCTCTTCAA 1980
Qy 3196 gaagaacttatgattctctctcctcagtgacacccaaagaaatggataaattagagaaa 3255
Db 1981 GAAGAATTTATCGATTCTCTCTCTCAGTGACAAACCAAGAAATGGATAAATAGAGAAA 2040
Qy 3256 accaagcagcgttacgaacacagacaccccttgaccttgccttgacagttgcagtgccaaat 3315
Db 2041 ACCAAGCAGCGCTTACGCAACAGAACCTTGACCTTGTGAAAGCTTGCAGGTGGCAAT 2100
Qy 3316 gtagagtcacaaagccttgagccaccattgagaagctcctgagcagtgagagcaagctg 3375
Db 2101 GGTAGGATCCAAAGCTTTGAGGCCACCATTTGAGAAAGCTCTCTGAGCAGTGAGACAACTG 2160
Qy 3376 aagcagggccatgcttaccttagaactggagcgtggtcgccctcctgcagcagcgtgagagag 3435
Db 2161 AAGCAGGCCATGCTTACCTTAGAAGTGGAGCGGTGGCCCTCTGTCAGACGCTGGAGGAG 2220
Qy 3436 ctgcggcggcggagcgacagcccgacccggcgctgagtcgagtcgacgcagcccgagccc 3495
Db 2221 CTGCGCGCGGCGAGCGACAGCCCGAGCGGAGCTGTGAGTCAGCAGCGCCGAGCCC 2280
Qy 3496 acggcgactgacagctctgcaggagagattgcaaacaccatccccactgtccaggcctt 3555
Db 2281 ACGGGCGACTGACAGCTCTGCAGGAGAGATTGCAACACCATCCACACTGTCCAGGCCTT 2340

RESULT 5

AK027355
LOCUS Homo sapiens cDNA FLJ14449 fis, clone HEMBB1001684.
DEFINITION AK027355
ACCESSION AK027355
VERSION AK027355.1 GI:14041975
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly body cDNA to mRNA, clone lib:HEMBB1 clone:HEMBB1001684.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS

Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Nakamura, Y., Nagahari, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2362)
 Isogai, T. and Otsuki, T.
 Direct Submission
 Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
 Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing;
 Research Association for Biotechnology; cDNA library construction,
 5'- & 3'-end one pass sequencing and clone selection; Helix
 Research Institute (supported by Japan Key Technology Center etc.)
 and Department of Virology, Institute of Medical Science,
 University of Tokyo,
 Location/Qualifiers
 1. .2362
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEMB1001684"
 /tissue_type="whole embryo, mainly body"
 /clone_lib="HEMB1"
 /dev_stage="embryo, 10 weeks"
 /note="cloning vector: pME18SFL3"
 40. .2064
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB5057.1"
 /db_xref="GI:14041976"
 /translation="MLKNKAKRSLTESLILSRGNKARGLOEHSIVDLSDLSSTL
 SNTSPVCEKEALPISPSFKLLGSSDLSDESHPPEAPLSPQAPRRANT
 LSHFPICEQEPQAPRSPGVQSKLMRYHSVSTETPHERKDFESKANHLGSGTPV
 KTRHRSQOIFLRVATPQKADSSRDYSELGELPPSPLEPCEDGPPGPPPE
 KRTSRELAKOKALQOILLRMEKENOKLOASENDLLRLKLDYBEITPCKEY
 TTWEMKLTSPGSKITDKFMKHAAGVGVPRHRRGEIWKFLAQFLHKQFPSPKQ
 PDVYPKELLKLTLSOHAILLDLGRTPHPVFSALGAGOLSLNILKAYSLDQ
 VYCGSLFVAGILLHMSFEAFKMLFMDGIRKQYRDMILLOIQVQLSRL
 HDYHRLYHNLHEHEIGPSLYAAPFLTFASQFPLGFVARVDFMIFLAGTEVIEKYA
 LSLGSHKPLLILOHENLETIVDFIKSTPLNGLVQMEKTINGVFMENIAKQLOAYEVE
 YHVLQELSDLSPLSDNRMDKLENTNSLRKONLDLLEQLQVANGRTSLRATIEKL
 LSESKLKQAMLTLELERSALLOTVLEELRRRSRSDREPECTQPEPTGD"

FEATURES

source

1804 atcgatgcccaggaaacctccaaacctgcccgggggtccccgggggttttcgcaaaagaaa 1863
 |||||
 361 ATCGAATGCCAGGAACCTCCAAACCTGCCGGGGGTCCCGGGGGTTTCGCAAGAGAAA 420
 |||||
 1864 cttatgaggtatcactcagtgagcacagacgcctcatgaacgaagagactttgaatcc 1923
 |||||
 421 CTTATGAGGTATACATCAGTCAGTCAGACACAGACGCCCTCATGACGAAGAAGGACTTTGAATCC 480
 |||||
 1924 aaagcaaaacattcttggtgattcttggtggaactctctgtgaagacccggagcattctctgg 1983
 |||||
 481 AAAGCAAAACCATCTTGGTGATTCTGGTGGACCTCCCTGTGAAGACCCGGAGGCATCTCTGG 540
 |||||
 1984 aggcacagatatctctcgagtagccaccccgcaagaagcggtgcgattcttccagcaga 2043
 |||||
 541 AGGCACAGATATTCCTCCGAGTAGCCACCCCGCAAGAGCGTGCATTCTTCCAGCAGA 600
 |||||
 2044 tatgaagattattcagagctgggagagcttccccccacgattctcttagaacagattgt 2103
 |||||
 601 TATGAAGATTATTCAGAGCTGGGAGAGCTTCCCCACGATCTCCTTTAGAACCATTTGT 660
 |||||
 2104 gaagaTggcccttggccccccacagaggaagaaagaaagacatctctgagctccga 2163
 |||||
 661 GAAGATGGGCCCTTTGGCCCCCACCACGAGAGAAAGAAAGACATCTCTGAGTCCGA 720
 |||||
 2164 gagctgtgcaaaagctattctcaacagatactctgcttagaattggaagagaaat 2223
 |||||
 721 GAGCTGTGCAAAAGGCTATTCTTCAACAGATCTGCTGTAGATTGGAAGAGGAANAAT 780
 |||||
 2224 cagaagctccaaagcctctgaaatgattgtgaaacagcgctgaagctcgattatgaa 2283
 |||||
 781 CAGAAGCTCAAGCCTCTGAAATGATTGTCTGAACAAGCGCTGAAGCTCGATTATGAA 840
 |||||
 2284 gaattactccctgtctttaaagaagtaactacagtgTgggaaagatgcttagcactcca 2343
 |||||
 841 GAAATTAATCTCTTAAAGAAAGTAACATACAGTGTGGGAAAAGATGCTTTAGCACTCCA 900
 |||||
 2344 gaaagataaaattaaattgattgacatggaaaataacacacgctgtgTggcgaagtgctg 2403
 |||||
 901 GGAAGATCAAAATTAAGTTTGACATGGAAAAATGACATCGGCTGTGGGCAAGGTGTG 960
 |||||
 2404 ccacgtcatccagaggtgaaatctggaatttctagctgagcaatttccacccttaaacac 2463
 |||||
 961 CCAGCTCATCACCGAGGTGAAATCTGAAATTTCTAGCTGAGCAATTCACCTTAAACAC 1020
 |||||
 2464 cagttccacagaacacagcagcgaaggatgtgccatacaaaagactcttaaaagcagctg 2523
 |||||
 1021 CAGTTTCCAGCAAAACAGCAGCAAAAGGATGTGCCATACAAAGAACTCTTTAAAGCAGCTG 1080
 |||||
 2524 actccacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 2583
 |||||
 1081 ACTTCCAGCAGCATCGGATCTTATTGACCTTGGGGAACCTTTCTTACACACCATAC 1140
 |||||
 2584 ttctctgcccagctgtggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2643
 |||||
 1141 TTCTCTGCCAGCTGTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1200
 |||||
 2644 ctctagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2703
 |||||
 1201 CTTCTAGACAGGAGTGGGATATTGCCAAGGCTCAGCTTTGTAGCAGGCATTTTGTCT 1260
 |||||
 2704 ctctcatgagtgaggaagagggcgtttaaaatgctcaagtctctgattgtgtgacatgggg 2763
 |||||
 1261 CTTTCAATGAGTCAGGAAGAGGCGTTTAAATGCTCAAGTTTCTGTGATTGTGACATGGG 1320
 |||||
 2764 ctgcgaaacagtatcggccagacagcagcagcagcagcagcagcagcagcagcagcagc 2823
 |||||
 1321 CTGCGAAACAGATATCGCCAGACATGATTTATTTTACAGATCCAGATGTACCAGCTCTCG 1380
 |||||
 2824 agattcttcaattaccacagacagcagcagcagcagcagcagcagcagcagcagcagcagc 2883
 |||||
 1381 AGGTTGCTTCATGATTACACAGAGACCTCTACAACTCCCTGGAGGACGACGAGATCGGC 1440
 |||||

BASE COUNT

ORIGIN

Query Match 59.4%; Score 2112; DB 9; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1444 agtgaattaccaccagtcgacgtgatttagctagatgctgaaaaaagcaag 1503
 Db |||||
 1 AGTGAATTACCCAGTGCCTCGATTAGGCTAGTATGCTGAAAAACAAGCAAG 60
 |||||
 QY 1504 agatcttaacagagctttgaaagattttgtccggggtaataagcagagcctg 1563
 Db 61 AGATCTTTAACAGAGCTTTTAGAAGATTTTGTCCGGGGTAATAAAGCCAGAGCCTG 120
 |||||
 QY 1564 caggaaacatccatcagtgTgattctggatagctccctctctagtaacattaagcaacc 1623
 Db 121 CAGGAACACTCCATCAGTGTGGATCTGGATAGCTCCCTGTCTAGTACATTAAAGTAACACC 180
 |||||
 QY 1624 agcaaaagacatctgtgtgaaagagccttgcctctctcctagagagcctcttaag 1683
 Db 181 AGCAAGAGGCATCTGTGTGTGAAAGAGGAGCCTTGCCCCATCTGTAGAGACTCCTTTAAG 240
 |||||
 QY 1684 ctctcggctcctcggagagcctgtccagtgactcgagagtagctatctcccagaagagcca 1743
 Db 241 CTCCTCGGCTCCTCGGAGGACCTGTCCAGTGACTCGGAGAGTCACTCCCGAGAGAGCCA 300
 |||||
 QY 1744 gctccgctgtgccccagcagcagccttcaggagcgagcagaacacccctgacttccctccc 1803
 Db 301 GCTCCGCTGTGCGCCAGCAGGCGCTTCAGGAGGCGAGCAAAACACCTTGAGTCACTTCCCC 360
 |||||

Qy 2444 agcaattccaccttaaacacaccaggtttcccccagcaaacagcagccaaagagtgccatata 2503
Db 609 AGCAGTACCACTGAAGCACCCTGTTTCCCTGCAMACAGCAGCCCAAGACACGCGCTACA 668
Qy 2504 aagaactcttaaacagcagctgactcccccagcagcagctgactcttatgacacctggcgaa 2563
Db 669 AAGAGCTTTAAAGCAGCTCACTCCCGAGCAGCAGCCATCTCATCGACCTTGGGGGAA 728
Qy 2564 cctttctccacacacccatactctctgcccagcttgagcagggagcagctatcgctttaca 2623
Db 729 CTTTCCAAACATACCTGCTACTCTGCTGCTAGGAGCAGGCGAGCTGCTCACTTACA 788
Qy 2624 acattttgaagcctactcaactcttagaccaggaagtggttatgccaaggtctcagct 2693
Db 789 ACATCTTGAAGCCTACTCGCTTCTAGACCAAGAGTGGGTATTTGCCAAGTCTCAGCT 848
Qy 2684 ttgtagcagcattttgcttctcatatgagtgaggaagcgcgtttaaaatgctcaagt 2743
Db 849 TTGTAGCAGGATCTCTGCTGCTCATATGGGTGAAGAGAGCGCTTAAACATGCTCAAGT 908
Qy 2744 ttctgattgttgacatgggctgcggaacacagtatcgcggccagacatgattattttacaga 2803
Db 909 TTTGATGTTTACATGGAGCTCGGAAACAATATCGGCGGACATGATTATTTTACAGA 968
Qy 2804 tccagatgtaccagctctcaggtgttctcatgattaccacagagacctctacaatcacc 2863
Db 969 TCCAGATGTACAGCTCTCGAGGCTGCTCCACGACTACACAGGAGGACCTCTACACACC 1028
Qy 2864 tggaggagcagagatcgcccccagcctctacgctgccccctgttctcaccatgttg 2923
Db 1029 TGGAGGACATGAGATGGCCCCAGCCCTCTATGCTGGCCCTGTTCTCACCCTGTTG 1088
Qy 2924 cctcacagttcccggtggattcgtagccagagcttttgatgatatttttttccaggaa 2983
Db 1089 CTTACAGTTCCCTCGGCTTCTGCGCAGAGCTTTTGATGATGATCTTTCTCAGGGAT 1148
Qy 2984 cagaggtcatttaagtggtcttaagctgttggtggaagccataagcccttgattctgc 3043
Db 1149 CAGAGGTCAATTTAAAGTGCTTTAAGTCTGCTGGGAAGCCATTAAGCCCTTGATCTGC 1208
Qy 3044 agcatgaaacctgaaacacatagttgactttataaaagcacgctaccacaccttgct 3103
Db 1209 AGCATGAAACCTAGAACCACTAGTTGACTTTTATAAAAGCACACTACCCCAACCTGGCT 1268
Qy 3104 tggtagagtgaaagacacacatcaggtatttgaaatggacatcgctaaacagttac 3163
Db 1269 TGGTGCAGATGGAGAAGACTATCAGTCAGGTGTTTGAAGAGGAGCTTATCGATTCTCTCTCA 1328
Qy 3164 aagcttatgagttgagtagcagctcctcaagaaagacattatcgattcctcctctca 3223
Db 1329 AAGCTTATGAAGTTGAGTACCACGTGCTTCAGGAAGAGCTTATCGATTCTCTCTCTCA 1388
Qy 3224 gtgacaaacaaagtgatgataattagagaacacacagcagcttacgcaaacagaacc 3283
Db 1389 CTGACAAACAAAGATGGACAACTAGAGAGACCAACAGCAGCTTACGCAACAGAAC 1448
Qy 3284 ttgacctctctgaaagtgagtgagtggaatgtaggatccaaagccttgagggccacca 3343
Db 1449 TTGACCTCTTGAACAGCTCGAGGTGCGCAATGGAGGATCCAGAGCCTGGAAGCCACTG 1508
Qy 3344 ttgagaagcctctgagcagtagagcaagctgaagcagggccatgcttaccttagaactgg 3403
Db 1509 TTGAGAAGCTCCTGACTAGCGAGAGCAAGCTGAAGAGGCGACCTGGCCTTGGAGTGG 1568
Qy 3404 agcgttcgcccctctgagcagcgttgagggagctgcgagcggcagcagcagcagcagc 3463
Db 1569 AGAGGTAGCCCTCTGTCGAGACGCTGGAGCAGCTCCGGAGGACACAGCGCAGAGTGGCA 1628
Qy 3464 accgggagcctgagtcagcagcagcccgagccacggcgagcagcagcagcagcagc 3511
Db 1629 GCCAGGAGTCTGACCCCNACACTGCCCAAGCCCTCAGGTGATGATGCTG 1676

RESULT 7

BC004675
LOCUS
DEFINITION Mus musculus, clone IMAGE:3500261, mRNA, partial cds.
ACCESSION BC004675
VERSION BC004675.1 GI:14919386
KEYWORDS house mouse.
SOURCE
ORGANISM
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NTH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalona@bcm.tmc.edu
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 9 Row: m Column: 4.

FEATURES

source
Location/Qualifiers
1..3051
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3500261"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCL_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1..1485
/codon_start=1
/product="Unknown (protein for IMAGE:3500261)"
/protein_id="AA04675.1"
/db_xref="GI:14919387"
/translation="ACDSPRYEDYSELGPPRSPLEPVCEGDPFGPQVEKKRTSR ELRLMKAILQIILLRMEKQKQASNDLLNKLKLDYDEITCLKEVTVWVK MLSTPGRSKIKFDMKVKHSAVGQSPRRHGEIWKFLAEQHLKHPSPQKPDVY KELLKLTQSOHAILDLGRTFPTHPVPSAQLGAGLSLYNLKAYSLLOQEVGCG LSFVAGILLHMESEERFKMLFLMEDMGLRKQYRPMIILQIOMYOLSLRLHDYHRD LYNLHEHEIGPSLIAFWLTFAFQPLGFVARVDFMFLQAGSEVIFKVALSLGS HKPLILQHENLETIVDFIKNTLPNLQVMEKTIQSFEMDIARQLOPIYEVTHVQGE ELIESPSLDNRMEKLEKNTSSRLKQNLDLLEQLQVANARISLEATVEKLLTSEK LKQAALTLEVERALLQMVLELRQSRPSTPEPDCPTQLEPTGD"

CDS

BASE COUNT 817 a 715 c 756 g 763 t
ORIGIN

Query Match 33.2%; Score 1179.6; DB 10; Length 3051;
Best Local Similarity 87.0%; Pred. No. 3.5e-247;
Matches 1296; Conservative 0; Mismatches 194; Indels 0; Gaps 0;

Qy 2023 gctgcgattcttccagcagatagaagattattcagagctggagagcttccccacaga 2082

Db 1 GCTTGTGACTCCCGCAGCAGATATGAAGATTATTCGAGCTGGGAGAGCTCCCTCCACGC 60

Qy 2083 tctccttagaacacagtttgtgaagatggcgcccttggccccccaccagagaaagaaa 2142

Db 61 TCCCTTTAGAACCGGTGTGTAGGAGCGGCCATTTGGCCCACTACAGGAGAAAGAGG 120
QY 2143 agacatctctgtatcgagagctgtgtgcaaaagcttatttcaacagatactctgtg 2202
Db 121 AAGACGTCACGCGAGCTTCGAGAGCTGTGGAAGAGGCCATCTTGACGAGATCCCTGCTG 180
QY 2203 ctLagaatgagagaaatcagaagctccagagctctgaaatgatttctgtaacaag 2262
Db 181 CTCAGATGGAGAGAGAGATCAGAGCTTACAAGCCTCTGAAACCGATTTGCTGGAACAA 240
QY 2263 cgctgaagctgattatgaagaaattactccctgtctttaaagaagttaactacagtggtg 2322
Db 241 CGCCTCAAGCTTGACATGATGAATCACTCCGCTGCTTAAACAAGTCACTACAGGTGG 300
QY 2323 gaaagatgcttagcactccaggaagatcaaaaatttaagtttgatcagtgaaataatgcac 2382
Db 301 GAAAGATGCTTAGCACTCCAGAGATCCAAATTAAGTTTGACATGGAAGAAAGTGCAC 360
QY 2383 tcgctgtgtggcaaggtgtgcacatcatccagaggtgaaatctggaatttctagct 2442
Db 361 TCAGCTGTGGGCAAGGTGTGCCAGCTCATCCAGAGGTGAGATCTGGAAATTTCTAGCT 420
QY 2443 gagcaattccacccttaaacaccagtttccagcaaacagcagcgaaggtgtgcatac 2502
Db 421 GAGCAGTTCACCTTAACACACCATTTCTAGTAAACAGCAGCCAAAGGACGTGCCCTAC 480
QY 2503 aagaactcttaagacagctgatttccagcagcatgcgatttcttattgacctggcgga 2562
Db 481 AAAGAGCTCTGAAGAGCTGACCTCGCAGCAGCAGCCATTTCTATCGACCTCGGGCGA 540
QY 2563 accttctctacacacatacttctctgccagcttggagcagcagcatctgcctttac 2622
Db 541 ACCTTTCCACACATCCATACCTTCTCTGCCAGCTTGGAGCAGGTGACGTGTCACATTTAC 600
QY 2623 aacatttgaagcctactactctctagaccaggaagtggtatattgcaaggtctcagc 2682
Db 601 AACATTCTGAAGCCCTACTCGCTCTGGACGAGGTTGGATACTGCCAAGTCTCAGC 660
QY 2683 ttgtgacaggaattttgtcttctctatagtgaggaagagcgttttaaaatgctcaag 2742
Db 661 TTTGTGCGAGGCAATTTGCTTCTCATAGTAGTGAAGAGAGGCGTTCAAGATGCTCAAG 720
QY 2743 ttctgatgttgacatgggctgcggaacagatctcggcagacatgattattttacag 2802
Db 721 TTCTGTATGTTGACATGGGCTCGCGAAACATATCGCCACACATGATTTATTTGCAG 780
QY 2803 atccagatgtaccagctctctogaggtgtctctatgattacacagagacctctacaatcac 2862
Db 781 ATCCAGATGTACCAGCTGTACCGGCTCTCCACGATTACCCAGACCTCTACAACCCAC 840
QY 2863 ctgagagacagagatggcccccagcctctacgtctgccccctgttctcaccatgttt 2922
Db 841 CTGGAAGACGACGAGATCGGCCACGCTCTAGCGGGCTCCCTGGTTCTCACCGGTTC 900
QY 2923 gcttcacagtcccgctgggattcgtagccagagcttttgatgatattttcttcaggga 2982
Db 901 GCCTCAGATTCCCTCGGCTTTGTAGCCAGAGTCTTTGATATGATCTCTCTCAGGGA 960
QY 2983 acagaggtcatatttaagtggttttaagctgtgtgggaagccataagcccttgattctg 3042
Db 961 TCAGAGGTATATTTAAAGTAGCTTTAAAGTCTTTTGGGAGCCATAAGCCCTTGATTTCA 1020
QY 3043 cagcatgaaacactagaaccatagttgactttataaaagacagctacccaacttggc 3102
Db 1021 CAGCATGAGAACCTGGAACCATCGTGGAGCTTCATAAAGAACACACTCCCCCAACCTGGGC 1080
QY 3103 ttggttcagatggaagaccatcaatcaggtatttgaaatggacatcgctaaacagtta 3162
Db 1081 CTGGTCAGATGGAGAGACCATCAGTCAGGTGTTTGAGATGGACATCGCCACAGCAGCTC 1140
QY 3163 caagcttatgaagttgagctaccagctctctcaagaagaacttatcgattctctctctc 3222

Db 1141 CAGCCCTATGAGGTGAGTACCAGTGCTCCAGGAGGAGCTTATTGAGTCTCGCCTCTC 1200
QY 3223 agtgacaaacaaagaatgataaattagagaaacaaacagcagcttagcgaacaaagaac 3282
Db 1201 AGTGACAACCAAGAATTTGGAGAAATTTGGAAACCAACAGCAGCTTGCCCAACAGAAC 1260
QY 3283 cttgactcctctgaacagttgcaggtggcaaatggtagatccaaagccttgaggccaac 3342
Db 1261 CTTGACTCTCTGGAGCAGTTGCAGGTGGCAATGCTAGGATCCAAAGCCTTTGAAGCCACG 1320
QY 3343 attgagaagctctcctgagcagctgagagcaagctgaaagcagcagcttaccttagaacty 3402
Db 1321 GTAGAGAAACTTCTTACCAGCAGAGAGTAAGCTGAAGCAGGCTGCCTGACCTGGAGTG 1380
QY 3403 gagcgtcggccctgtgcagacggtggagagagctcggcggcgagcagcagccagc 3462
Db 1381 GAGCGTCCGCCCTGTCTGAGATGTTGGAGGAGCTCGGAGGCAAGCCCGGCCCGCAGC 1440
QY 3463 gaccggagcctgagtcgacgagccgagccagccagcggcgagcagctgacagct 3512
Db 1441 ACTCCAGAGCCAGCTGCACCCAGCTGGAGCCACAGGCGATTTGACCGCT 1490

RESULT 8

AK057182 1788 bp mRNA linear PRI 31-OCT-2001
Homo sapiens cDNA FLJ32620 fis, clone STOMA2000386, highly similar
to Bos taurus mRNA for lycein.
AK057182
ACCESSION AK057182.1 GI:16552778
VERSION oligo capping; fis (full insert sequence).
KEYWORDS Homo sapiens stomach cDNA to mRNA, clone_lib:STOMA2
SOURCE clone:STOMA2000386.
Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (sites)
Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T.,
Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R.,
Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H.,
Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E.,
Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S.,
Terashima,Y., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T.,
Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project

TITLE

Unpublished

JOURNAL

2 (bases 1 to 1788)

REFERENCE

Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES

source

1..1788
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="STOMA2000386"
/tissue_type="stomach"
/clone_lib="STOMA2"
/note="cloning vector: pME18SFL3"

BASE COUNT	481 a	412 c	434 g	461 t
ORIGIN				
Query Match	28.1%; Score 998.4; DB 9; Length 1788;			
Best Local Similarity	99.9%; Pred. No. 1.3e-207;			
Matches 999; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
Qy	2556 tggcggaaccttctctacacaccatcttctctgcccagcttgagcagcagacgctatc 2615			
Db	553 TGGCGGAACCTTCTCTACACACCCTACTCTCTGCCAGCTTGAGCAGGACGACTATC 612			
Qy	2616 gctttacacatttgaaagcctactcaactctagacacagaaagtggatattccaaag 2675			
Db	613 GCTTTACAACTTTTGAAGGCTTACTCTTCTAGACAGGAAGTGGATATTCCTCAAGG 672			
Qy	2676 tctcagcttttagcagcattttgctctcttcattatgagtgagaaagcggtttaaaat 2735			
Db	673 TCTCAGCTTTTAGCAGCACTTTTGTCTTCTCATATGATGAGTGAAGAGGCGTTTAAAT 732			
Qy	2736 gctcaagttctgatgttgatcattggtgctcggaacacagtatcgccagacatgattat 2795			
Db	733 GCTCAAGTTCTGTGATGTTGACATGGGGCTCGGAAACAGTATCGCCAGACATGATTAT 792			
Qy	2796 ttacagatccagatgtaccagctctcgaggttcttcattatgattaccacagacacctta 2855			
Db	793 TTTACAGATCCAGATGACAGCTCTCGAGGTTGCTTCATGATTTACACAGAGACCTCTA 852			
Qy	2856 caataccttgagagcagcagatcgcccccagcctctacgctgcccctgggttcctcac 2915			
Db	853 CAATCACCTGGAGGACGACGAGATCGGCCCCAGCCTCTACGCTGCCCTGTTCTCCAC 912			
Qy	2916 catgtttccctcaagttcccgctgggattcgttagcagagctcttgatgattttct 2975			
Db	913 CATGTTTCCCTCACAGTTCCTCGGCTGGGATTCGTAGCCAGAGCTCTTGATATGATTTTCT 972			
Qy	2976 tcagggaacagaggtccattattaaagtggctttaaagctgttggaagccataagccctt 3035			
Db	973 TCAGGGAACAGAGGTCATATTTAAAGTGCGCTTAAAGTCTGTGGGAAGCCATAGACCTT 1032			
Qy	3036 gattctgcagcatgaaacacctagaaaccatagttgactttataaaaagcacgtaccocaa 3095			
Db	1033 GATTCTGCAGCATGAAACACCTAGAAACCATAGTTGACTTTATATAAAGACAGCTACCCAA 1092			
Qy	3096 ccttgcttggtacagatggaaagaccatcaatcaggattttgaaatggacatcgctaa 3155			
Db	1093 CTTTGGCTTGTACAGATGGAAGAGACCATCAATCAGGTATTTGAAATGGACATCGCTAA 1152			
Qy	3156 acagttacaagcttatgaagttgagtagccagctctcctcaagaagaacttatcgattcctc 3215			
Db	1153 ACAGTTACAAGCTTATGAAGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1212			
Qy	3216 tctctcagtgacacacaaagattgattgattgattgattgattgattgattgattgattg 3275			
Db	1213 TCTCTCAGTGACAAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272			
Qy	3276 acagaaccttgacctctccttgagctgagctgagctgagctgagctgagctgagctgag 3335			
Db	1273 ACAGAACCTTGACCTCTCTGAAACAGTTCAGGTGGCAATGGTGGATGCAAGGCTTGA 1332			
Qy	3336 ggcacacattgagaagctcctgagcagtgagcagcagcagcagcagcagcagcagcagc 3395			
Db	1333 GGCCACCAATGAGAAGCTCTTGACGATGACAGCAAGCTGAAGCAGGCGCATGCTTACCT 1392			
Qy	3396 agaactgagcagctgagcagctgagcagcagcagcagcagcagcagcagcagcagcagc 3455			
Db	1393 AGAAGTGGAGCGGTGCGGCCCTGCTGACAGCGTGGAGAGCTGCGGCGGCGGAGCGCAGA 1452			
Qy	3456 gccacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3515			
Db	1453 GCCCAGCGACCGGAGCGCTGAGTGCACGACGCGCGGCGGCGGCGGCGGCGGCGGCGG 1512			
Qy	3516 caggagagattgcaac 3555			

Db 1513 CAGGAGAGATTGCAACACCATCCACACTGTCACAGGCTT 1552

RESULT 9

AB011175

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FEATURES

Location/Qualifiers

1..5922

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HG1488b"

/sex="male"

/tissue_type="brain"

/clone_lib="pBluescriptII SK plus"

348..4247

/gene="KIAA0603"

348..4247

/gene="KIAA0603"

/codon_start=1

/product="KIAA0603 protein"

/protein_id="BA25529.1"

/db_xref="GI:3043730"

/translation="MEPPSCIQDEPPHPLEPEPGVSAQPGPKPSDKRFLRLTVGGS

CLDRTTLPMLPMAEIRRSQKPEAGCGCAPAAREVILVSAPFLURCPAPGAGAS

GGTSPATQNPVAFIPEHKAQHSIRIHNHDLTYPAYLIKAPDDPSOMACHVFR

ATDPSQVDPVSIISROLKAKEDAKPKDNEDAFYNSQKFLVYCGKVTVTHKRAP

SSLIDCMKESLHEOORLKTQGEORGPDEDLADLEVVPVPGSGCLPEADGTOT

HLGPGAGSQPALTSSRVCFPERILESGDEQEQFRSCSVTVGVORRHESQSKSO

PRRHASAPSHVQPSDEKRNLTQVGFEEINLTSTKSVLEKFNKDLSSCSQOI

KHYDHGFIRESPEGLSQICYVQCASSELVDEYMLTKQAFSTAAALQSAKTI

KLCACPMHSILKLCERIEGLYPRAKLVIOHLSLTDEQADIFERVQKMPVSDO

EENELVTLHLRCLCEAKQTHVIGEPSTISNATIPENATSSGRFKDLILKNKARS

LTSLENI FSGANRMRGLSVDSFERSNLSASEKDYSPGDSPGTPPASPPSSAWQ

TFPEEDSDSPQFRRAHTSHPSSTTKRLNLQDGRAQGVRSPLLRQSSSECSNLS

VRRMYKESNSSSLPSLHTSPSAPSTFKSLFSYQNSGRLSFQIENEIRQDTASES

SDGGRKRTSTCSNESLTVGTSVTPRRISWRQRIFLVASPNKSPSAMQDDGLD

RNELLPLSLPTMEEPPLVIFSGDEDDPEIKERKSKELSLRWAKIRCDMDIHT

EKENKLEGASRDELOSKVKLDYEEVGAQKEVLTWDKLLNKLAKIRCDMDIHT

LLEGVPSRSGEITWFLAQYLRHLRHLNPKQOPDPSIYKELLKQLTAQHAILVDLG

RTPTPTFTYVQLGPGQLSNLLKAYSLDLKVEGYCGGIFSVAGVLLHMSQAPE

MLKFLMYDLGFRKOYRDMMSLIQMYQLSKLLHDYHRLDNLHLENEISFSLYAAPW

FLTLFASQSLGVARVDIIFLQGTVEVIFKVALSLSSQETLMECESFENIVEFLK

NTLPDMNTSEMKIITOVFEMDIKOLHAYEVEYHVLQDELQESSYCESEETLEKLE

RANSOLKRONMDLLEKLOVAHTKIOALENLENLLTRETMMKSLIRTLEQEKMAQKT

VEQLRKLPPADLANCLDLLRLDNCNPNKAKIGNKP"

1624 a 1408 c 1432 g 1458 t

BASE COUNT

ORIGIN

	Query Match	19.2%	Score 682.2;	DB 9;	Length 5922;
	Best Local Similarity	56.6%;	Pred. No. 1.9e-138;		
	Matches 1585; Conservative	0;	Mismatches 1028; Indels 186; Gaps 10;		
Qy	823	agcgacacaataatgtgcagcccaacagatatcaggaataatcgaactatgctcttcacg	882		
Db	1368	AGGCACCCAGTCAGTCCAGCCTCGGACTCGGNAGAGACAGACGCATGCTCTCCAG	1427		
Qy	883	attgccagctcgaagttaacctcatcagtccttgacacccccaaaaaatagcattggagaaa	942		
Db	1428	GTTGGGGATTGTGAGATTAACACTTATCATAGTCAGACACATAATCATAGTTGTGTAGAAAAG	1487		
Qy	943	aatttaaggagatatccttttgcctcagggccatcacagctgagaccatttgggttt	1002		
Db	1488	AATTTTAAGAATATCTCTCTTGTTCTCAGGGTATAAAGCAATGTGGATCACATTTGGCTTT	1547		
Qy	1003	atctgtcggggagctctccggaggctggcgcttcatatttgtctgaagcttttcagtg	1062		
Db	1548	ATCTGCCGGAGTCTCCAGAGCCTGGACTTAGCCAGTATATTGTTATGTATTCCAGTGT	1607		
Qy	1063	acaaatgaggctctgtgttgatgaataattgatgacctgaaacaggccottcaagtgccc	1122		
Db	1608	GCCAGCGAATCTCTGGTTGTATGAGTAATGTGACTCTGAAACAGGCCCTTCAGTAGCGCG	1667		
Qy	1123	gcagtgcagcagacagctaaagcgccagccagctgtgtgagggctgcccctgcacaaagc	1182		
Db	1668	GCTGCCCTGCAGATGCCAAGCCAGATTAACTGTGTGAGGCCTGCCGATGCACCT	1727		
Qy	1183	ctgcacagctctgtgagaggatagaggggaatgaattcttccaaacaaaactagaactg	1242		
Db	1728	TTCATTAAGCTCTGTGAAGGATTGAAGGTCTCTACCACCAAGAGCCAAGCTGTGTGATA	1787		
Qy	1243	caaagcacctgcagcaattaaccaatcagagcagcgactatttttgaagaggttcag	1302		
Db	1788	CAGAGGCATCTCATCTACTGACAGATTAATGAGCAAGCTGACATCTTTGAAGAAGATTGAG	1847		
Qy	1303	aaattgagccgagaaatlgagcagcgagagaatgaattgattttctttcttgagatgt	1362		
Db	1848	AAAAATGAAGCCAGTCAGTGACAGCAGGAAGAAATGAATGTGATTTACACCTGAGGCAG	1907		
Qy	1363	ttaatgaagagaanaacagaagaacacatccatatgtggggagatgaagcagacatcgcag	1422		
Db	1908	CTGTGTGAAGCCACAGCAAGACACATGTGCACATCTGGGG-----AAGGCCCTTCT	1958		
Qy	1423	atggcagcagagaatatlggaagtgaattacaccaccagtcgcactcgatttaggctagat	1482		
Db	1959	ACTATTTCAAATAGTACAATCCCAGAAAATGCAACAACAGCAGTGGGAAGGTTCAAACTGCAC	2018		
Qy	1483	atgctgaaaaaacagaagaagatctttaacagagcttttagaagatattttgtccogg	1542		
Db	2019	ATTCTGAAAAATAAAGCTTAAGAGATCTCTTAACCTAGCTCCCTGGAAAAATCTCTCAAGG	2078		
Qy	1543	gg---taataaagccagagccctgcaggaacactccaactg-----1581			
Db	2079	GGAGCTAACAAATGAGAGAGTCGGCTTGGAGTGTGGACAGTTTTGAACGGTCCACACAGT	2138		
Qy	1582	-----1581			
Db	2139	CTTGCTTCAGAGAAGGACTACTCACAGGGGATTTCTCCACCAGGACACCCGACGGTCC	2198		
Qy	1582	-----gtggatcgtgtagatcctcctgtcctagt	1608		
Db	2199	CCACGGTCTCAGCTTGGCAACGTTTTCCGAAAGAGGATTCGGATCTCCCGCAGGTTTGA	2258		
Qy	1609	acattaaagtaaccacagcaaaagaccatctgtgtgtaaaggagggccttggccatctct	1668		
Db	2259	AGACGGGCACACACGCTTCAGCCACCCACCTTCAAGCACAAAGAGAAAGCTGAATTTGCAG	2318		
Qy	1669	gagagctcctttaagctcctcctgcctcgc-----gagg	1702		
Db	2319	GATGGGAGGGCTCAGGGTGTGGTTCCCTCTGCTGAGCGACAGCTCCAGTGAACAGTGC	2378		

Qy	1703	acctgtccagtactcgaggagtcattccccagaagagccagctccgctgtcgccccagc	1762
Db	2379	AGCAATCTTTTCGTGAGTTCGACGCATGTACAAGGAGAGTAATTTCTTCTCCAGTCTTCCA	2438
Qy	1763	aggccttcaggagcgagcaaacacacctgagtcacctcccacatgcaatgccagaaacctc	1822
Db	2439	AGTCTTCACACTTCTCTCTGCCCCCTCTCTTCACTGCCCCCTCTTCTCCTGAAAAGCTTT	2498
Qy	1823	cacaaacctgcgccgggggtc---cccggggggtttccaaaggaaccttatgaggtatcaact	1879
Db	2499	TACCAGAAATTCAGGTAGACTGTCCCCACAGTATGAATGAATCAGACAGACACTGCT	2558
Qy	1880	cagtgcagacagagcgcctcatgcaacgaaaggacatttgaatccaaagacaaacctattg	1939
Db	2559	TCAGAATCAAGTGATGAGAAAGGAGAAAGGAGCTCATCTACCTGCAGCAATGAGTGCC	2618
Qy	1940	gtgaattc-ggtgggactcctgtgaagaccggaggacattcctggaggcacagatatc	1998
Db	2619	CTAAGTGTGGGAGGAACTCTGTGTCACCTCTGCCCGGATCTCTGTGGCGACGCGATTTTC	2678
Qy	1999	ctccgagtagcccccgcgagaaaggcgtgcgattcttcacgcagatatgaagattattca	2058
Db	2679	CTCAGGGTTCCTTCCATGACAAATCTCCCTCAGCAATGCCAACAGATGGATTG	2738
Qy	2059	gagctgggagagcttcccccaogatctctcttagaacccagtttgtgaagatggggccctt-	2117
Db	2739	GACAGGAACGAGCTGCTGGCAGTGTCCCCCTCTCTCCAACATGAGGAGGAGAAACCGCTG	2798
Qy	2118	-----tgcccccccccacagaggaaagaaagacacatctcgt	2154
Db	2799	GTTATATCTGTCTGGGAGGATGACCCAGAAAAAGATTGAAGAAAGAAAGAAATCAAAA	2858
Qy	2155	gagctccgagagctgtggcaaaaggctattcttcacagatactcgtcttgaagtggag	2214
Db	2859	GAACTGAGGAGCTGTGGGAAAGACTATACCAACAAATCTTGTACTTTCGAATGAA	2918
Qy	2215	aaggaaatcagaagctcca---agcctctgaaatgatgttgcgaacaagcgctgaag	2271
Db	2919	AAAGAAAACCAAGAACTTGAAGGAGCAAGCAGAGATGAATCCACAGTCCAGAAAAAGTTAAA	2978
Qy	2272	ctcgattatagaataactcctctctttaaagaagtaactacagtgctgggaaagatg	2331
Db	2979	TTAGACTATGAAGAAGTTGGTGATGTCAAGAAGAGGTCTTAATNACTTTGGGATPAGAAG	3038
Qy	2332	cttagcactccaggaagatcaaaaaaagtttgatcagtaaaatactcagctcgctgt	2391
Db	3039	TTGTTAA--ACTGCAGACTAAATCAGATCTGATATGGAAGATATTATACATCTTCTT	3095
Qy	2392	gggcaaggtgtgccagtcataccogaggtgaaatcttgaaaaattctagtgcagcaattc	2451
Db	3096	AAAGAAGGAGTTCCAAAAGTCGACGAGGAGAAATTTGGCAGTTTCTGGCTTTACAGTAC	3155
Qy	2452	caccttaaacacacagtttcccagcaaacagcagccaaaggatgtgcatacaagaagact	2511
Db	3156	CGACTCAGACACAGATTGGCTTAATAAACCAACAGCCTCTGTACATATCTATPAGGAATCT	3215
Qy	2512	ttaagcagctgcattcccagcagatgcgatctcttattgaccttggcggaacctttcct	2571
Db	3216	TTGAGCAGCTCACTGCTCAGCAGCATGCGATTCCTCGTGGATTTAGGAAGGACGCTTCTCT	3275
Qy	2572	acacaccaactctctgtgccagcttggagcagcagcagctatcgtcttacaacatttg	2631
Db	3276	ACTCACCTTACTTTTCAGTACAGCTTTGGCCGAGACAGCTGTCACTCTTTTAACTCCTCG	3335
Qy	2632	aaggcctaactcctctagaccaggaagtgggatatttgcgaaggtctcagcttttgtagca	2691
Db	3336	AAAGCCATTCTTTGTGTGGACAAGAAGTGGGATACTGTCTAGGGGATCAGCTTTGTGGCT	3395
Qy	2692	ggcattttgtcttctcatatgagtgcaggaagcggttttaaaatgctcaagtttctgatg	2751
Db	3396	GGAGTCTGCTTCTGACATCTGAGTGAAGAGCAAGCCTTTGAAATGCTGAAATTCCTCATG	3455
Qy	2752	tttagacatggggtgcqqaacacagtatcgccagacacagattattttacagatccagatg	2811


```
CC * 26385 31223: contig of 4839 bp in length
CC * 31224 31323: gap of 100 bp
CC * 31324 36014: contig of 4691 bp in length
CC * 36015 36114: gap of 100 bp
CC * 36115 41939: contig of 5825 bp in length
CC * 41940 42039: gap of 100 bp
CC * 42040 48122: contig of 6083 bp in length
CC * 48123 48222: gap of 100 bp
CC * 48223 52983: contig of 4761 bp in length
CC * 52984 53083: gap of 100 bp
CC * 53084 58094: contig of 5011 bp in length
CC * 58095 58194: gap of 100 bp
CC * 58195 64606: contig of 6412 bp in length
CC * 64607 64706: gap of 100 bp
CC * 64707 71079: contig of 6373 bp in length
CC * 71080 71179: gap of 100 bp
CC * 71180 79886: contig of 8707 bp in length
CC * 79887 79986: gap of 100 bp
CC * 79987 86155: contig of 6169 bp in length
CC * 86156 86255: gap of 100 bp
CC * 86256 95370: contig of 9115 bp in length
CC * 95371 95470: gap of 100 bp
CC * 95471 103712: contig of 8242 bp in length
CC * 103713 103812: gap of 100 bp
CC * 103813 113565: contig of 9753 bp in length
CC * 113566 113665: gap of 100 bp
CC * 113666 123867: contig of 10202 bp in length
CC * 123868 123967: gap of 100 bp
CC * 123968 132296: contig of 8329 bp in length
CC * 132297 132396: gap of 100 bp
CC * 132397 141110: contig of 8714 bp in length
CC * 141111 141210: gap of 100 bp
CC * 141211 161622: contig of 20412 bp in length
CC * 161623 161723: gap of 100 bp
CC * 161723 189811: contig of 28089 bp in length.
```

Key Location/Qualifiers

```
FT source 1. .189811 /Chromosome="4"
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /map="4"
FT /clone="RP11-392K14"
FT /clone_lib="RP11 Human Male BAC"
FT misc_feature 1. .451
FT misc_feature 552. .2663
FT misc_feature 2764. .4124
FT misc_feature /note="assembly_fragment"
FT misc_feature /note="assembly_fragment"
FT misc_feature 4225. .6767
FT misc_feature /note="assembly_fragment"
FT misc_feature 6868. .8750
FT misc_feature /note="assembly_fragment"
FT misc_feature 8851. .11272
FT misc_feature /note="assembly_fragment"
FT misc_feature 11373. .13411
FT misc_feature /note="assembly_fragment"
FT misc_feature 13512. .16962
FT misc_feature /note="assembly_fragment"
FT misc_feature 17063. .22544
FT misc_feature /note="assembly_fragment"
FT misc_feature 22645. .26284
FT misc_feature /note="assembly_fragment"
FT misc_feature 26385. .31223
FT misc_feature /note="assembly_fragment"
FT misc_feature 31324. .36014
FT misc_feature /note="assembly_fragment"
FT misc_feature 36115. .41939
FT misc_feature /note="assembly_fragment"
FT misc_feature 42040. .48122
FT misc_feature /note="assembly_fragment"
```

```
FT misc_feature 48223. .52983
FT /note="assembly_fragment clone_end:SP6 vector_side:left"
FT misc_feature 53084. .58094
FT /note="assembly_fragment"
FT misc_feature 58195. .64606
FT /note="assembly_fragment"
FT misc_feature 64707. .71079
FT /note="assembly_fragment"
FT misc_feature 71180. .79886
FT /note="assembly_fragment"
FT misc_feature 79987. .86155
FT /note="assembly_fragment"
FT misc_feature 86256. .95370
FT /note="assembly_fragment"
FT misc_feature 95471. .103712
FT /note="assembly_fragment"
FT misc_feature 103813. .113565
FT /note="assembly_fragment"
FT misc_feature 113666. .123867
FT /note="assembly_fragment"
FT misc_feature 123968. .132296
FT /note="assembly_fragment"
FT misc_feature 132397. .141110
FT /note="assembly_fragment"
FT misc_feature 141211. .161622
FT /note="assembly_fragment"
FT misc_feature 161723. .189811
FT /note="assembly_fragment"
FT XX
SQ Sequence 189811 BP; 52455 A; 40839 C; 40321 G; 53485 T; 2711 other;
```

```
Query Match 13.1%; Score 465.8; DB 30; Length 189811;
Best Local Similarity 99.6%; Pred. No. 5.2e-91;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 416 aagtcctgagatcatcagctccatccgtcagcgagggaagatcgcccgagagagc 475
DB 97313 AGGTGCTGAGATCATCAGCTCCATCCGTCAGCGGGGAGATCGCCCGCAGGAGAGC 97254
QY 476 tgcactgcccgtccgaggttcgacgacacgttttccaaagattcgaggtcttcttgcg 535
DB 97253 TGCACGTGCCGTCCGAGTTCGACGACACGTTTCCAAAGATTTCGAGGTGCTCTTCTGCG 97194
QY 536 gcgcgtgacgttgcgacaaagaagctccgcgcctgacgacgagtgatcgatcgaga 595
DB 97193 GCGCGCTGACGTGGCGGCACAAAGAGGCTCCGCGCCCTGTATCGACGAGTGCATCGAGA 97134
QY 596 agttcaatcacgtcagcgagcgggggtccgagagcccccgcccccaaccccgccatg 655
DB 97133 AGTTCAATCAGTCAGCGCAGCGCGGGGTCCGAGAGCCCCCGCCCAACCCGCCCATG 97074
QY 656 ccgcgccccacaggagccaggagcctgtgcgagggcccatgagcaagtccttctccagc 715
DB 97073 CCGCGCCACAGGAGCCAGGAGCCTGTGCGAGGCCCATGCGCAAGTCTCTTCCGAGC 97014
QY 716 ccgcgctgcgtcgtgcttttaggaagagctgagagatgggggctcccgagagcagcg 775
DB 97013 CCGGCTGCGCTCGCTGGCTTTAGGAAGGAGCTGGGGGCTCCGAAGCAGCG 96954
QY 776 gcttctcagctccttcgagagagcgacatgagaccacctcattagcgacacaata 835
DB 96953 GCTTCTTCAGCTCCTTCGAGGAGGAGCGACATTTGAGAACCCACCTCATTTAGCGGACAAATA 96894
QY 836 ttgtcagccccacagatatcgaggaaaaatcgaaactatgctcttcacgat 884
DB 96893 TTGTGACGCCACAGATATCGAGGAAAATCGAACTATGCTCTTTCACGGT 96845
```

```
RESULT 11
AC021106 195108 bp DNA linear PRI 03-JUL-2001
LOCUS AC021106
DEFINITION Homo sapiens clone Rp11-177C12, complete sequence.
```



```
/chromosome="4"
/clone="RP11-392K14"
misc_feature 1..25746
  /note="assembly_name:Contig16"
misc_feature 25847..65131
  /note="assembly_name:Contig17"
misc_feature 65232..105028
  /note="assembly_name:Contig18"
misc_feature 105129..159307
  /note="assembly_name:Contig19"
misc_feature 159408..208318
  /note="assembly_name:Contig20"
BASE COUNT 59195 a 45373 c 44728 g 58621 t 401 others
ORIGIN

Query Match      13.1%; Score 465.8; DB 2; Length 208318;
Best Local Similarity 99.6%; Pred. No. 5.2e-91;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 416 aatgctgagatcatcagctccatcgcgtcaggcggggaagatcgccggcaggagagc 475
  |||
Db 129553 AGTGGCTGAGATCATCAGCTCCATCCGTCAGCGGGGGAAGATCGCCCGGACAGGAGC 129494
  |||
QY 476 tgcactgcccgtccgagttcgagcacagcttttcccaagaagttcgcgtctcttcgcg 535
  |||
Db 129493 TGCACATGCGCGTCCGAGTTCGACACACGCTTTTCCCAAGAAGTTCCGAGGTGCTCTTCGCG 129434
  |||
QY 536 gccgctgacggtggtggcgacaaaggctccgcccgcctgatcgacgagtgcatcgaga 595
  |||
Db 129433 GCCGCTGAGCGGTGGCGCACAGAAGGCTCCGCGCGCCCTGATCGACGAGTGATCGAGA 129374
  |||
QY 596 agttcaatacgtcagcgcagcgcgtgggtccgagagcccccccccacccgccccatg 655
  |||
Db 129373 AGTTCAATCAGCTCAGCGCAGCGGGGTCCGAGAGCCCCCGCCCAACCCGCCCATG 129314
  |||
QY 656 ccgcgccacagagccagagacgtgtgcgagggcccatgagcagctctctccagc 715
  |||
Db 129313 CCGCGCCACAGAGCCAGAGACCTGTGCGCAGGCCCATGCGCAAGTCTTCTCCAGC 129254
  |||
QY 716 ccggctgcgctgcgtggccttttagaaggagctgcagatggggccctccgaagcagcg 775
  |||
Db 129253 CCGGCTGCGCTCGCTGGCCCTTTAGGAAGAGCTGCAGGATGGGGCTCCGAAGCAGCG 129194
  |||
QY 776 gcttttcagctcttcgagagagagacattagaaacctcattagcgacacaata 835
  |||
Db 129193 GCTTCTCAGCTCCTTCGAGGAGAGCGACATTCAGAACCCATTCATAGCGGACACAATA 129134
  |||
QY 836 ttgtcagccacagatattcgagagaaatcgaaactatgctcttcacgat 884
  |||
Db 129133 TTGTGCGAGCCACAGATATCGAGGAAATCGAACTATGCTCTTCACGGT 129085
  |||

RESULT 13
AL390202.04/c
WPCOMMENT
Sequence split into 10 fragments LOCUS AL390202 Accession AL390202
Fragment Name Begin End
AL390202_00 1 110000
AL390202_01 100001 210000
AL390202_02 200001 310000
AL390202_03 300001 410000
AL390202_04 400001 510000
AL390202_05 500001 610000
AL390202_06 600001 710000
AL390202_07 700001 810000
AL390202_08 800001 910000
AL390202_09 900001 988176
Continuation (5 of 10) of AL390202 from base 400001 (AL390202 Homo sapiens chromosome 20)

Query Match      11.9%; Score 424.2; DB 2; Length 110000;
Best Local Similarity 96.0%; Pred. No. 6.1e-82;

Matches 435; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 1 atggaaccaataaacattcacagcaaggaaacattctcttctactaacagaggtctcgtggat 60
  |||
Db 10365 ATGGAACCAATAAACATTCACAGCAAGAAACATCTCTCTTAAAGAGTCTCGGTGGAT 10306
  |||
QY 61 ttggcctcagctggtgggtccctgcctgtgcatctcctgacccaccatgcccatgctg 120
  |||
Db 10305 TTGGCCTCAGCTGTTGGGTCCCTGCTGTGTCATTCCTTGACCATGCCCATGCTG 10246
  |||
QY 121 ccttgggttggctgaggtgcgaagactcagcaggcagtcaccaccaggaaggaaacctgta 180
  |||
Db 10245 CCTGGGTTGGCTGAGGTGCGAAGACTCAGCAGGACAGTCCACAGGAAAGGAACTGTA 10186
  |||
QY 181 accaagcaagtcggttgcgtttcaacctctggactgaagatgaaacctgagccagg 240
  |||
Db 10185 ACCAAGCAAGTCGGGCTTTGCGFTTCACCTCTGGACTGAGATGTGAACCTGAGCCAGG 10126
  |||
QY 241 agaagtcacagtggtggtccctgatctattccagcatctttgagtgcaagcctcagcgt 300
  |||
Db 10125 AGAAGTCACAGTGGGATCCCTGATCTATTCCAGCATCTTTGAGTGCAAGCCTCAGCGT 10066
  |||
QY 301 gtccacaaactgattcacacagtcagtcacccaagtatgacccaagttaattgctgtgtgtaaggaa 360
  |||
Db 10065 GTTCACAAACTGATTCACAACTCATGACCCAAAGTTACTTTGCTGTCTGTATTAAGGA 10006
  |||
QY 361 gacgctgtccacggcagagtatctgctatgttctcaagccgagatatacaacaaagt 420
  |||
Db 10005 GACGCTGTCCACGGCAGAGTATCTGTATGTGTTCAAGCCGATGATCAACAAAGTA 9946
  |||
QY 421 cctgagatcatcagctccatccgtcagcgcg 453
  |||
Db 9945 AGTGAGATGGAGATCCAAAGACTAAGGTGTGG 9913
  |||

RESULT 14
DMU50542
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
fruit fly strain-Oregon R.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 3214)
  Zhang,S.D., Kassisi,J., Olde,B., Mellerick,D.M. and Odenwald,W.F.
  Pollux, a novel Drosophila adhesion molecule, belongs to a family
  of proteins expressed in plants, yeast, nematodes, and man
  Unpublished
REFERENCE
2 (bases 1 to 3214)
  Zhang,S.D.
  Direct Submission
  Submitted (04-MAR-1996) Shang-Ding Zhang, Neurogenetics, LNC,
  NINDS, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1..3214
  /organism="Drosophila melanogaster"
  /strain="Oregon R"
  /db_xref="taxon:7227"
  /chromosome="3"
  /map="83c"
  /tissue_type="trachea; CNS"
1..3214
  /gene="plx"
130..2328
  /gene="plx"
  /function="adhesion molecule"
  /codon_start=1
  /product="pollux"
  /protein_id="AAB02200.1"
```

```
/db_xref="GI:1373163"
/translation="MRPAKRGKRDAAELRELWRTAIRTQIMLRNMTENAMLAQON
ENELKRLDYEELVPCDKQLIERWEOIERNSTQIGNKKDKPVLGHAIIRTVGPRSK
GDVMTFLAEQHSMTAPVDTRKRFNFNTPYHMLKGLHTEHQHAFIDLGRTPFNHQFY
KDPGLGQLSFLNLKAYSILDPDELGYCOGLGFCVLLHLHCDSEANSFOLLKHLMFRR
NMRTKYLPMKMLFOLVSLRSLVKDHLPLDLYVLDNDVDSPTLYAAAPWILTVSSOF
PLGVARVDFLLFLESSDVIKFAIALLSVHKQOLLANDNFEEIMDYLKTVPKMEHT
CMEQIMKLVSMDIGKQLAEYNEYNVLQBELITNNHLEMLNREKTONQHLQOLQF
AQSIALETTRSSQQAQITLQVQSLELTQTLGRYVQLVEHPDLDELPEVNER
MLQDLDDLRORRPIETERIGKSVSNHSLGPKLVLELTERDELGSFKOKKELK
TPEFQELQOQQHRLPAGGQSSNVGSGSPTPPSRPNRLIDNASARTVMQVQKOKKEK
LPHVDKRVANIKSLPEVDVGVTPLSPPTASNSGSGISFRMGVYRTTPPALSLAQ
RQSGVATTTAPCPHMEEVAPATTMAVPOEDVEEPQWPLPSMVGSDVNVREKGT
OLKSIRPVHHRAPILPLGGVQHPSTPEAVRVPVVELAPPAATATGRS"
BASE COUNT      862 a 798 c 774 g 780 t
ORIGIN

Query Match      8 6%; Score 304.6; DB 3; Length 3214;
Best Local Similarity 56.0%; Pred. No. 6.1e-56;
Matches 621; Conservative 0; Mismatches 479; Indels 9; Gaps 2;

Qy 2155 gagctccgagagctggaagaagcttcttcacagatactgctcttagaagag 2214
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 169 GAGCTGAGGGAACCTGCGGCACCTGCCATTAGCGAGACTATAATCTGAACCGCATGGAG 228

Qy 2215 aaggaaaaatcagaagctccaaagcctctctgaaatgatttgtaacaagcgctgaagctc 2274
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 229 ACGGAGAACGCCATTGTGCGAGCAGCGGAGAGATAGAGACGAGCTAAAGCGCATTAACATG 288

Qy 2275 gattatgaagaataactcctctgtcttaagaagtaactacagtgatggaaagatgctt 2334
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 289 GACTACGAGGAGATTGTACCCCTGCGACAGACGAGCTAATCGAGCGATGGGAGCAGATCAAT 348

Qy 2335 ---agcaccctccggaagaacaaataagtttgatggaataatgcactcgctgtt 2391
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 349 GAGCGCAACTCACACAGATAGCAACAGAGGATCCCAAGTCTCGGCCACGCCATT 408

Qy 2392 gggcaagtggtccacgctcaccagagtggaatctctggaataattctagctgagcaatc 2451
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 409 CGCACCGGAGTACCGGATCGAAGCGCGGCGATGCTGGACCTTCCTGGCTGAGCAGCAT 468

Qy 2452 caccctaa-----acaccagtttcccagcaaacagcagcacaagatgtgcacacaaa 2505
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 469 TCCATGACACGCGACCGGTGGACACAAAGCGATTCCCAACTTCAATACACCGGTATCAC 528

Qy 2506 gaactcttaaaagcagctgacttcccagcagcaatgcgattcttattgaccttggcgcaacc 2565
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 529 ATGCTGCTGAACACCTTACTGAGCATCAGCATCGGATTTTCAATTGATCTGGCGAGGACT 588

Qy 2566 ttctctacacacccatacttctctccagcttgtagcagcagcagctatcgctttacaac 2625
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 589 TTTCTTAATCACCACTGTACAAAGATCCGCTTGGCTCGCGCCACTTTCGCTGTTTAAAC 648

Qy 2626 atttgaagccctactcactcttagaccagaagtggaatattgcaaggtctcagcttt 2685
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 649 CTGCTGAAGGCGTACTCCATTCTAGATCCGGAATTAGGGTACTGCGCAGGCTCGGCTTC 708

Qy 2686 gtacgagcagcttctcttcttctatgatgtaggagagggcgttttaaaatgctcaagttt 2745
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 709 ATCTCGGGCGTCTTACTCTCTGATTCGATGAAGCAATTCATTCAACTGCTGAAGCAC 768

Qy 2746 ctgagtgattgacatgggctcggaagaacagtatcgccagacatgattattttacagatc 2805
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 769 CTAATGTTTTTCGCGGTAATATATGCGACGAAATACCTTCGCGGACATGAAAAAGTTTCAACTG 828

Qy 2806 cagatgtaccagctctcaggttctctatgattaccacagagacacctacaatcacctg 2865
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 829 CAGCTGTATCAGCTCTCCGATTTGTCAGAGGATATCTGCCAGATCTCTACGTGTGGCTC 888

Qy 2866 gaggagcacgagatcgccccagcctctacgtcgccctgttctcaccatgtttgcc 2925
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 889 GATCAGACGATGCTCTCGCCCACTTTGTTATGCGGGCTCCATGATGATCCTCACCCTCTTTAGC 948
```

```
Qy 2926 toacagttcccgctgggagtttagtagccagagctctttgatgatgatttttttcagggaaca 2985
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 949 TCACAGTTTCGCGTGGGATTCGTGGCCCGCGCTCTTTTGATCTGCTCTTCTTGGAACTCCTCC 1008

Qy 2986 gaggtccattataagtgacttaagtctgttgggaagccataagcccttgattcttgcag 3045
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1009 GATGTAATCTTTAACTTTTGCCATTCCTTCCTCCGTGTCACCAAGCAGCACTACTTGCC 1068

Qy 3046 catgaaacctagaaacaccatagttgactttataaaagcagcctaccccaaccttgcttg 3105
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1069 AAGGATAACTTCGAAGACATATTGACTATCTTGAAGACCGTGGTCCAAAGATGGAGCAC 1128

Qy 3106 gtacagatggaaaaagacacatcaatcaggtatttgaaaaggacatcgctaaaacagtata 3165
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1129 ACGTGCATGGAGCAAAATTTAAGCTGTCTTTAGCATGGACATTTGGAAGCAGCTTGCC 1188

Qy 3166 gcttatgaagttgtagtaccagctcttcaagaagaacttatcgattctctctctcagt 3225
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1189 GAATACATATGGAGTACAATGTGCTTCAGGAGGAGATTACCACCACCTAACCATCACCTA 1248

Qy 3226 gacacccaagaatggataaattagagaa 3254
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1249 GAAATGCTTAACCGGGAGAGACGCGAGAA 1277
```

RESULT 15

```
LOCUS DME17919 4746 bp mRNA linear INV 18-NOV-1998
DEFINITION Drosophila melanogaster mRNA for pollux protein.
ACCESSION Y17919
VERSION Y17919.1 GI:3893102
KEYWORDS pollux gene.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 4746)
AUTHORS Xu,X.Z., Wes,P.D., Chen,H., Li,H.S., Yu,M., Morgan,S., Liu,Y. and
            Montell,C.
            Retinal targets for calmodulin include proteins implicated in
            synaptic transmission
            J. Biol. Chem. 273 (47), 31297-31307 (1998)
REFERENCE 2 (bases 1 to 4746)
AUTHORS Montell,C.
            Direct Submission
            Submitted (11-AUG-1998) C. Montell, The Johns Hopkins University
            School of Medicine, Department of Biological Chemistry, 725 N.
            Wolfe Street, Baltimore, MD 21205-2185, USA
```

FEATURES

```
source
    1..4746
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
    1..3316
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /map="83C"
        /tissue_type="retinal"
        /dev_stage="adult"
        452..4591
            /gene="pollux"
            /gene="pollux"
            /note="calmodulin binding protein"
            /codon_start=1
            /product="Pollux protein"
            /protein_id="CAA76939.1"
            /db_xref="GI:3893103"
            /translation="MAELMHQMRDPAHTLGGSVGSIPTQLIGGGGGSHGNSGNALNGI
            HAYPATNLKMEAMRNAOHDTSANPVSCKMKASKSYTHGLSSSGTWNITPTSAQSN
            LSLADISPNTHEFFVNYVKIKVSKRPNTFIDDALPKFAYDAQRLRLQNKKM
```

BASE COUNT ORIGIN	1211 a	1342 c	1261 g	932 t																	
	SLSEGGVTEAKPSSLSKSHDLKBEDEEEQKHGDDSDSQAKPLVLQQLTGAEE	GAAPRLDNEKNKSEPKRLLRGQQLIEGHKEGSDGSPSAANSGLQKAPNVVNNQ	TPPRDQGVGTASASAGSPLQHPNYAMDNTPKQDRSASQGIPIPYEQNTMTFL	VGRCDDILSPDRKQVLKDYKDVASCVHGQSKSLDFHGIGICELRNDNGYIGYVFCQ	SEHVCDDIVAAIAQAEPDTCNAEKQKQDQITFSCCHCPMLWYHKLNDVDEGLSEKKTQA	LILURRTELSDQETVWAKFCGSEKTSNPVAEQNOFLMMLLRAHCESSQOORHVHDTA	ENESPELNOYLGGSITFKAKRSILTNFSEONLKRKPSKODIAPVPSHNLDIRBGSAP	LQGTQSPPEGRPSRNSVNGASPSCKPTAEQLKSPMDI FIKVGNSPKEAZETHQGSWRQA	ILNSVQSPGLOSEVPTPELSPMRKPAKPKRDKDALEEDLRLMRTA LQTIIMLNMRETE	NAMLOARONELKRLKLDYEEVPCDDKQIIEWQI IERNSTQIGNKKDKPVLGHAI	RTPNHQYKDPGLGQSLFNLNKAYSTIDPELPGVCGQLGICGVLLLHDCDEANSEFO	LAKHLERNMRTKYLPDMKKFOLQYLSRLYKDHLPOLYVWLDQNDYSPTYLAPAM	LILTVSSQPPPLGFVARVFDLLFLESSDVFKPAIALLSVHKQQLLAKNFEEIMDLK	TVYPMKMTCEQIMKLVSMDIGLOLAENYVNIQLQEITTTNHNHLEMLNREKTON	QHLEUQLOFAQSIAQIETRRSQQAETITLQTSQVSELTITQTLRGYVQLVHEPND	LELPENVRMLQOJLDLDRQRPIETERRIGQSVNSHLGSPPLKVLBELTEHRDELG	SPOKQDEKLPFEUQKROOHLNGGQSSNVGSGSPPTPPSRNRLNDNASARTV	MOYKLBELKLPEHVDKRVANIKSPLEVDSCVGTPLSPPTASNSGSSITFSRMGYRTT	PPALSPLAQOQSQGVALTATPACPHHMEEVAPATTMAMPQOEDEEBEQPHPLPSWVGDD	VNRYFKGTQLKSIKRPVHHMRAIPLGCVQHPSPSTEPFAVRVAPVVELAPPAATATTGR	S*

Query Match	8.6%	Score 304.6	DB 3	Length 4746
Best Local Similarity	56.0%	Pred. No. 6.4e-50		
Matches 621; Conservative	0	Mismatches 479	Indels	Gaps
Qy 2155	gagctccgagagctgtggcaaaagctattcttcaacagatactcgtcttagaatggag	2214		
Db 2432	GAGCTGAGGGAACGTGGCGCACTGCCATTAGGCAGACTATAATGCTGAACCGCATGGAG	2491		
Qy 2215	aaggaaatcagaagctcccaagcctctgaaatgatttgcgtgaacaagcgctgaaagctc	2274		
Db 2492	ACGGAGAAACGCCATGTTGTCAGACACGGCAGAAATGAGAACGAGCTAAAGCGCATTTAAACTG	2551		
Qy 2275	gattcataaagaaattactccctcgtctaaagaagtaactacagctgtgggaaaaagatgctt	2334		
Db 2552	GACTACGAGGAGATTGTACCTTCGCACACGACAGCTAATCGAGCGATGGGACGACATCAATT	2611		
Qy 2335	--agcactccaggaagatcaaaaaattaagtgttgacatggaaaaaaacgactcggctgtt	2391		
Db 2612	GAGCGCAACTCCACACAGATAGGCAACAAGAAAGATGCCAAAGTCTGGGCACGCGCATT	2671		
Qy 2392	ggcaagtgctgcacatcatcaccgagtgaaatctggaattcttagcttgagcaattc	2451		
Db 2672	CGCACCGGAGTACC CGGATTCGACGCGCGCATGCTTGGACCTTCTTGGCTTGACGACGAT	2731		
Qy 2452	caccttaa-----acaccagtttcccagcaaaacagcagcaaaagatgtgccatacaaa	2505		
Db 2732	TCCATGAACACGGCACCGGTGGACACAAGGATTCGCCAACTTCAATACACCGTATCAC	2791		
Qy 2506	gaactctaaagcagctgaacttccagcagcatcgcatttcttagaccttggcgcaacc	2565		
Db 2792	ATGCTGCTGAAACACCTTACTAGCATCAGATCGCATTTTCATTTGATCTGGGCAGGACT	2851		
Qy 2566	ttctctacacaccataactctctgcccagcttggagcagcagactatcgctttacaac	2625		
Db 2852	TTTCCTAATCACCAAGTCTTACAAGGATCCGCTTGGTCTCGGCCAGTTGTGCTGTTTAAAC	2911		
Qy 2626	atttgaaggcctactcactcttagaccaggaagtgggatatattgccaaaggtctcagcttt	2685		
Db 2912	CTGCTGAAGGCGGTACTCCATTCTAGATCCGGAATTAGGGTACTGCCAGGGTCTGGGCTTC	2971		
Qy 2686	gtacaggcatcttgcttcttcataatgagtgaagaagcgcttttaaataagctcaagttt	2745		
Db 2972	ATCTGGCGCTGTACTCTGCATGCGATGCAACCAATTCAATTCACTGCTGAAGCAC	3031		
Qy 2746	ctgattgttgacatggggctcgaaaaacagtatcggccagacatgattattttacagatc	2805		
Db 3032	CTAATGTTTCCCGTAATATATCGACAGAAATACCTGCGCGACATGAAGAACTTTCAACTG	3091		

Search completed: September 9, 2002, 19:16:17
Job time: 17490 sec

Qy	2806	cagatgtaccagctctcgcgaggttgctctcatgattaccacagagacctctacaatacacctg	2805
Db	3092	CAGCTGTATCAGCTCTCCCGATTTCGTCAGGAGTCATCTGCCAGATCTCTACGTGTGGCTC	3151
Qy	2866	gaggagcacagatcgcgcccgacgtcctcagctgcgccctgcttccctcaccatgcttgc	2925
Db	3152	GATCAGAACCATGTCTCGCCCACTTTGTTATGGGGCTCCATGGATCCTCACCGTCTTTAGC	3211
Qy	2926	tcacagttcccgctgggattctgagccagagctctttgatgatgattttcttcagggaaca	2985
Db	3212	TCACAGTTTCCGTGGATTCGTGGGCCGCGCTTTTGATCTGCTCTCTGCGAATCCCTCC	3271
Qy	2986	gaggctcatattaaagtggctttaagctctgtgggaagccataaagcccttgattcttcgag	3045
Db	3272	GATGTAATCTTTAAGTTTGGCATTTGCTCTCGCTGCACAGCAGCACTACTGCCC	3331
Qy	3046	catgaaaacctagaaaccaagttgactctttaaaaaagcacgtaccaccaacctggcttg	3105
Db	3332	AAGGATAACTTCGAAGAGATTATGGACTATCTGAAGACCGTGGTGCCAAAGATGGAGCAC	3391
Qy	3106	gtacagatgaaagaccatcaatcaggctatttgaaatggacatcgctaaacagttacaa	3165
Db	3392	ACGTGCATGGAGCAAAATTATGAAGTGGCTTTTACCATGGACATTTGGAAAGCAGCTGCC	3451
Qy	3166	gcttataagttgagtaccacgctccttcagaagaactctatcgattctctctctcctcagt	3225
Db	3452	GAATACAAATGTGGNGTACAATGTGCTTTACGAGGAGAGATTACCACCACCTAACCATCATCCTA	3511
Qy	3226	gacaaccaaagaatgataaaattagaaa	3254
Db	3512	GAATGCTTTAACCGGGGAGAGACGCAGAA	3540

